



TGP/8/1 Draft 16

ORIGINAL: englisch

DATUM: 1. Oktober 2010

**INTERNATIONALER VERBAND ZUM SCHUTZ VON PFLANZENZÜCHTUNGEN**  
GENF

**ENTWURF**

Verbundenes Dokument  
zur  
Allgemeinen Einführung zur Prüfung auf  
Unterscheidbarkeit, Homogenität und Beständigkeit und zur  
Erarbeitung harmonisierter Beschreibungen von neuen Pflanzensorten (Dokument TG/1/3)

**DOKUMENT TGP/8**  
**PRÜFUNGSANLAGE UND VERFAHREN FÜR DIE PRÜFUNG DER**  
**UNTERSCHIEDBARKEIT, DER HOMOGENITÄT UND DER BESTÄNDIGKEIT**

*vom Verbandsbüro erstelltes Dokument*

*vom Rat während seiner vierundvierzigsten ordentlichen Tagung  
am 21. Oktober 2010 in Genf zu prüfen*

<u>INHALTSVERZEICHNIS</u>	<u>SEITE</u>
EINLEITUNG.....	5
TEIL I: DUS-PRÜFUNGSANLAGE UND DATENANALYSE.....	6
1. DUS-PRÜFUNGSANLAGE.....	6
1.1 Einleitung.....	6
1.2 Wachstumsperioden.....	7
1.2.1 Einleitung.....	7
1.2.2 Unabhängige Wachstumsperioden.....	7
1.3 Prüfungsort.....	8
1.3.1 Zweck.....	8
a) Verkürzung der gesamten Prüfungsperiode.....	8
b) Reserveprüfung.....	8
c) Verschiedene agroklimatische Bedingungen.....	9
1.3.2 Verwendung von Informationen aus mehreren Prüfungsorten.....	9
a) Zusätzliche Prüfungen.....	9
b) DUS aufgrund von Daten, die für dieselben Merkmale an verschiedenen Prüfungsorten erfaßt wurden.....	9
1.4 Bedingungen für die Durchführung der Prüfung.....	10
1.5 Prüfungsanlage.....	10
1.5.1 Einleitung.....	10
1.5.2 Anzahl Pflanzen im Versuch.....	10
1.5.3 Prüfungsanlage.....	10
1.5.3.1 Einleitung.....	10
a) <i>Einzelparzellen</i> .....	11
b) <i>Einzelpflanzen und Parzellen in Reihen</i> .....	11
c) <i>Wiederholte Parzellen</i> .....	11
1.5.3.2 Einzelparzellen.....	14
1.5.3.3 Wiederholte Parzellen (statistische Analyse).....	14
1.5.3.3.1 Einleitung.....	14
1.5.3.3.2 Wiederholte Parzellen für die statistische Analyse der Daten von Einzelpflanzen..	14
1.5.3.3.3 Randomisierung.....	14
1.5.3.3.4 Randomisierte unvollständige Blockanlagen.....	18
1.5.3.3.5 Anlage für Paarvergleiche zwischen bestimmten Sorten.....	19
1.5.3.3.6 Weitere statistische Aspekte der Prüfungsanlage.....	20
1.5.3.3.6.1 Einleitung.....	20
1.5.3.3.6.2 Die Hypothesenprüfung.....	20
1.5.3.3.6.3 Bestimmung der optimalen Stichprobengröße.....	22
1.5.3.3.7 Elemente des Anbauversuchs bei Einsatz der statistischen Analyse.....	23
1.5.3.3.7.1 Einleitung.....	23
1.5.3.3.7.2 Parzellen und Blöcke.....	24
1.5.3.3.7.3 Zuordnung der Sorten zu den Parzellen.....	25
1.5.3.3.7.4 Größe, Form und Anordnung.....	25
1.5.3.3.7.5 Unabhängigkeit der Parzellen.....	26
1.5.3.3.7.6 Anordnung der Pflanzen innerhalb der Parzelle / Typ der Parzelle für die Beobachtung.....	26
1.5.3.4 Randomisierte Blindprüfungen.....	26
1.6 Änderung der Verfahren.....	26
2. VALIDIERUNG DER DATEN UND ANNAHMEN.....	27
2.1 Einleitung.....	27
2.2 Validierung der Daten.....	27
2.3 Annahmen für die statistische Analyse und Validierung dieser Annahmen.....	29
2.3.1 Annahmen für die statistische Analyse, die mit der Varianzanalyse verbunden sind.....	30
2.3.1.1 Einleitung.....	30
2.3.1.2 Unabhängige Erfassungen.....	30
2.3.1.3 Varianzhomogenität.....	30
2.3.1.4 Normalverteilte Erfassungen.....	32
2.3.1.5 Additivität der Block- und Sorteneffekte.....	32
2.3.2 Validierung der Annahmen für die statistische Analyse.....	33
2.3.2.1 Einleitung.....	33
2.3.2.2 Datendurchsicht.....	34
2.3.2.3 Erstellung von Abbildungen.....	34

3.	WAHL DER STATISTISCHEN VERFAHREN FÜR DIE PRÜFUNG DER UNTERSCHIEDBARKEIT .....	39
	3.1 Einleitung .....	39
	3.2 Statistische Verfahren bei zwei oder mehreren unabhängigen Wachstumsperioden.....	39
	3.2.1 Einleitung.....	39
	3.3 Zusammenfassung der ausgewählten statistischen Verfahren für die Prüfung der Unterscheidbarkeit.....	41
	3.4 Voraussetzungen für die statistischen Verfahren für die Prüfung der Unterscheidbarkeit .....	42
TEIL II: AUSGEWÄHLTE VERFAHREN FÜR DIE DUS-PRÜFUNG .....		43
1.	DIE GAIA-METHODIK .....	43
	1.1 Gründe für die Summierung und Gewichtung erfaßter Unterschiede .....	43
	1.2 Berechnung des phänotypischen GAIA-Abstandes.....	44
	1.3 Detaillierte Informationen über die GAIA-Methodik.....	44
	1.3.1 Gewichtung der Merkmale.....	44
	1.3.2 Beispiele für Anwendungen.....	46
	1.3.2.1 Bestimmung der „Unterscheidbarkeit plus“ .....	46
	1.3.2.2 Weitere Beispiele für Anwendungen .....	47
	1.3.3 Berechnung des phänotypischen GAIA-Abstandes.....	48
	1.3.4 GAIA-Software.....	48
	1.3.5 Beispiel mit Daten von <i>Zea mays</i> .....	50
	1.3.5.1 Einleitung.....	50
	1.3.5.2 Analyse der Noten.....	50
	1.3.5.3 Elektrophoreseanalyse .....	52
	1.3.5.4 Analyse der Messungen .....	54
	1.3.5.5 Messungen und Skala 1 bis 9 beim selben Merkmal .....	56
	1.3.6 Beispiel für eine GAIA-Bildschirmkopie.....	57
2.	ELTERNFORMEL VON HYBRIDSORTEN.....	60
	2.1 Einleitung .....	60
	2.2 Voraussetzungen des Verfahrens.....	60
	2.3 Prüfung der Eigenständigkeit einer neuen Elternlinie .....	60
	2.4 Überprüfung der Formel.....	62
	2.5 Homogenität und Beständigkeit der Elternlinien.....	62
	2.6 Beschreibung der Hybride.....	62
3.	DAS KOMBINIERTES UNTERSCHIEDBARKEITSKRITERIUM ÜBER MEHRERE JAHRE (COYD) .....	63
	3.1 Zusammenfassung der Voraussetzungen für die Anwendung des Verfahrens .....	63
	3.2 Zusammenfassung .....	63
	3.3 Einleitung .....	64
	3.4 Die COYD-Methode .....	64
	3.5 Anwendung von COYD .....	65
	3.6 Anpassung von COYD an besondere Gegebenheiten .....	66
	3.6.1 Unterschiede zwischen Jahren im Ausprägungsbereich eines Merkmals .....	66
	3.6.2 Geringe Anzahl von Sorten in den Prüfungen: Langzeit-COYD .....	66
	3.6.3 Deutliche Änderungen von Jahr zu Jahr bei einem einzelnen Sortenmerkmal .....	67
	3.7 Umsetzung von COYD.....	67
	3.8 Quellenangaben .....	68
	3.9 Statistische Verfahren der COYD .....	71
	3.9.1 Varianzanalyse.....	71
	3.9.2 Modifizierte kombinierte Regressionsanalyse (MJRA) .....	71
	3.9.3 Vergleich von COYD mit anderen Kriterien .....	72
	3.10 Die COYD-Software .....	73
	3.11 Schemata für die Anwendung von COYD .....	80
4.	DAS 2X1 %-VERFAHREN.....	83
	4.1 Voraussetzungen für die Anwendung des Verfahrens.....	83
	4.2 Das 2x1 %-Kriterium (-verfahren) .....	83
5.	DER CHI-QUADRAT-TEST NACH PEARSON MIT KONTINGENZ-TABELLEN.....	84
6.	EXAKTER FISHER-TEST .....	88
	6.1 Prüfung der Unterscheidbarkeit.....	88
7.	<b>MATCH-ANSATZ</b> .....	91
	7.1 Voraussetzungen für die Anwendung der Methode.....	91
	7.2 Match-Methode .....	91

8.	VERFAHREN FÜR DIE PRÜFUNG DER HOMOGENITÄT ANHAND VON ABWEICHERN .....	92
8.1	Fester Populationsstandard .....	92
8.1.1	Einleitung .....	92
8.1.2	Anwendung des Verfahrens zur Prüfung der Homogenität bei einer Pflanzenart .....	92
8.1.3	Bei der Entscheidung über die Anwendung des Verfahrens zu berücksichtigende Aspekte ..	94
8.1.4	Beispiele .....	95
8.1.5	Einführung in die Tabellen und Abbildungen .....	100
8.1.6	Verfahren für eine einzige Prüfung .....	101
8.1.7	Methode für mehr als eine einzige Prüfung (Prüfungsjahr) .....	102
8.1.7.1	Einleitung .....	102
8.1.7.2	Kombinierte Prüfung .....	102
8.1.7.3	Zweiphasenprüfung .....	102
8.1.7.4	Sequenzprüfungen .....	104
8.1.8	Anmerkung zum Ausgleich der Fehler vom Typ I und vom Typ II .....	104
8.1.9	Definition der statistischen Begriffe und Symbole .....	104
8.1.10	Tabellen und Abbildungen .....	106
9.	DAS KOMBINIERTES HOMOGENITÄTSKRITERIUM ÜBER MEHRERE JAHRE (COYU) .....	113
9.1	Zusammenfassung der Voraussetzungen für die Anwendung des Verfahrens .....	113
9.2	Zusammenfassung .....	113
9.3	Einleitung .....	114
9.4	Das COYU-Kriterium .....	115
9.5	Anwendung von COYU .....	115
9.6	Mathematische Einzelheiten .....	116
9.7	Frühzeitige Entscheidungen für eine Dreijahresprüfung .....	119
9.8	Beispiel für COYU-Berechnungen .....	121
9.9	Umsetzung von COYU .....	123
9.10	COYU-Software .....	123
9.10.1	DUST-Computerprogramm .....	123
9.11	Schemata für die Anwendung von COYU .....	127
10.	PRÜFUNG DER HOMOGENITÄT ANHAND VON VERFAHREN DER RELATIVEN VARIANZ .....	130
10.1	Anwendung des Verfahrens der relativen Varianz .....	130
10.2	Schwellenwert für verschiedene Stichprobengrößen .....	130
10.3	Der Test der relativen Varianz in der Praxis .....	131
10.4	Beispiele für das Verfahren der relativen Varianz .....	131
10.5	Beziehung zwischen relativer Varianz und relativer Standardabweichung .....	132

## **EINLEITUNG**

Zweck dieses Dokuments ist es, Anleitung zur Prüfungsanlage und Datenanalyse zu geben und Informationen über bestimmte für die DUS-Prüfung angewandte Verfahren zu erteilen. Dieses Dokument ist wie folgt aufgebaut:

TEIL I: DUS-PRÜFUNGSANLAGE UND DATENANALYSE: gibt Anleitung zur Prüfungsanlage, Datenvalidierung und zu den für die statistische Analyse zu erfüllenden Annahmen.

TEIL II: VERFAHREN FÜR DIE DUS-PRÜFUNG: Gibt Einzelheiten zu bestimmten Verfahren an, auf die in den Dokumenten TGP/9 „Prüfung der Unterscheidbarkeit“ und TGP/10 „Prüfung der Homogenität“ verwiesen wird und für die weitere Anleitung angebracht erscheint.

Ein Überblick über die Teile des Prozesses der Prüfung der Unterscheidbarkeit, bei denen die Prüfungsanlage und die Verfahren, die in diesem Dokument behandelt werden, von Belang sind, ist in Dokument TGP/9 „Prüfung der Unterscheidbarkeit“, Abschnitt 1 „Einleitung“ zu finden.

## **TEIL I: DUS-PRÜFUNGSANLAGE UND DATENANALYSE**

### **1. DUS-PRÜFUNGSANLAGE**

#### **1.1 Einleitung**

1.1.1 Anleitung für die Durchführung der Prüfung wird in den Prüfungsrichtlinien gegeben, sofern solche verfügbar sind. Verschiedene Prüfungsrichtlinien wurden erstellt, und es werden laufend weitere ausgearbeitet. Eine aktuelle Liste ist in Dokument TGP/2, „Liste der von der UPOV angenommenen Prüfungsrichtlinien“ und auf der UPOV-Website ([http://www.upov.int/de/publications/tg\\_rom/](http://www.upov.int/de/publications/tg_rom/)) enthalten. Die UPOV empfiehlt jedoch folgendes Verfahren, um Anleitung zur Prüfung der Unterscheidbarkeit, der Homogenität und der Beständigkeit zu geben, wenn keine Prüfungsrichtlinien vorliegen.

#### Erfahrung anderer Verbandsmitglieder mit der DUS-Prüfung

1.1.2 Das Prüfungsamt wird gebeten, das Dokument TGP/5, „Erfahrung und Zusammenarbeit bei der DUS-Prüfung“ (<http://www.upov.int/de/publications/tgp/>) und die GENIE-Datenbank (<http://www.upov.int/genie/de/>) zu konsultieren, um festzustellen, ob andere Verbandsmitglieder über praktische Erfahrung bei der DUS-Prüfung verfügen.

1.1.3 Wenn solche Erfahrung vorhanden ist, werden die Sachverständigen ersucht, sich an die betreffenden Verbandsmitglieder zu wenden und gemäß den Grundsätzen in der Allgemeinen Einführung zu versuchen, ihre Prüfungsverfahren möglichst weitgehend zu harmonisieren. In einem nächsten Schritt werden die Verbandsmitglieder ersucht, die UPOV gemäß den in Dokument TGP/5 „Erfahrung und Zusammenarbeit bei der DUS-Prüfung“ angegebenen Maßnahmen über das Vorhandensein des harmonisierten Prüfungsverfahrens zu unterrichten oder gegebenenfalls zu empfehlen, daß die UPOV-Prüfungsrichtlinien für die betreffende Art erstellen soll.

#### DUS-Prüfungsverfahren für neue Arten oder Sortengruppierungen

1.1.4 Wenn in anderen Verbandsmitgliedern keine praktische Erfahrung mit der DUS-Prüfung für die betreffende Art oder Sortengruppierung vorhanden ist, müssen die Sachverständigen eigene Prüfungsverfahren entwickeln.

1.1.5 Die Ämter werden dazu angehalten, bei der Entwicklung solcher Prüfungsverfahren diese an die in der Allgemeinen Einführung (Dokument TG/1/3) dargelegten Grundsätze und an die Anleitung zur Erstellung von Prüfungsrichtlinien in Dokument TGP/7 „Erstellung von Prüfungsrichtlinien“ anzupassen. Weitere Anleitung ist in Dokument TGP/13 „Anleitung für neue Typen und Arten“ zu finden.

1.1.6 Das Prüfungsverfahren sollte gemäß den Anforderungen für die Prüfungsrichtlinien insoweit dokumentiert werden, als die Erfahrung und die Informationen dies erlauben.

1.1.7 Gemäß der Anleitung in der Allgemeinen Einführung und in Dokument TGP/7 befolgt dieser Abschnitt den Aufbau des Abschnitts 3 der UPOV-Prüfungsrichtlinien, „Durchführung der Prüfung“.

## **1.2 Wachstumsperioden<sup>1</sup>**

### **1.2.1 Einleitung**

1.2.1.1 Eine wichtige Überlegung bezüglich der Anbauversuche ist die Bestimmung der geeigneten Anzahl Wachstumsperioden. Diesbezüglich sagt das Dokument TGP/7, Anlage I, TG-Mustervorlage, Abschnitt 4.1.2, folgendes aus:

#### „4.1.2 Stabile Unterschiede

Die zwischen Sorten erfaßten Unterschiede können so deutlich sein, daß nicht mehr als eine Wachstumsperiode notwendig ist. Außerdem ist der Umwelteinfluß unter bestimmten Umständen nicht so stark, daß mehr als eine Wachstumsperiode erforderlich ist, um Gewißheit zu erlangen, daß die zwischen Sorten beobachteten Unterschiede hinreichend stabil sind. Ein Mittel zur Sicherstellung dessen, daß ein Unterschied bei einem Merkmal, das in einem Anbauversuch erfaßt wird, hinreichend stabil ist, ist die Prüfung des Merkmals in mindestens zwei unabhängigen Wachstumsperioden.“

1.2.1.2 Die UPOV-Prüfungsrichtlinien, sofern solche vorliegen, geben die empfohlene Anzahl Wachstumsperioden an. Bei der Formulierung der Empfehlung berücksichtigen die Sachverständigen, die die UPOV-Prüfungsrichtlinien ausarbeiten, Faktoren wie die Anzahl der im Anbauversuch zu vergleichenden Sorten, den Einfluß der Umwelt auf die Merkmalsausprägung und den Grad der Variation innerhalb der Sorten unter Berücksichtigung der Besonderheiten der Vermehrung der Sorte, beispielsweise ob sie eine vegetativ vermehrte, selbstbefruchtende, fremdbefruchtende oder Hybridsorte ist.

1.2.1.3 Wenn die UPOV keine individuellen Prüfungsrichtlinien für eine bestimmte Art oder sonstige(n) Gruppe(n) erstellt hat, sollte die Prüfung gemäß den insbesondere in der Allgemeinen Einführung festgelegten Grundsätzen, insbesondere den Empfehlungen in Abschnitt 9 „Durchführung der DUS-Prüfung bei Fehlen von Prüfungsrichtlinien“, erfolgen (vergleiche Absätze 1.1.1 bis 1.1.7).

### **1.2.2 Unabhängige Wachstumsperioden**

1.2.2.1 Wie in Abschnitt 1.2.1.1 angegeben, ist die Prüfung des Merkmals in mindestens zwei unabhängigen Wachstumsperioden ein Mittel zur Sicherstellung dessen, daß ein Unterschied bei einem Merkmal, das in einem Anbauversuch erfaßt wird, hinreichend stabil ist.

1.2.2.2 In der Regel beruht die Beurteilung der Unabhängigkeit auf der Erfahrung der Sachverständigen.

1.2.2.3 Wenn ein Merkmal in einem Anbauversuch in zwei unabhängigen Wachstumsperioden erfaßt wird, wird es in der Regel in zwei getrennten Anpflanzungen oder Aussaaten beobachtet. Bei einigen mehrjährigen Pflanzen wie Obstbäumen nehmen die Wachstumsperioden die Form eines in zwei aufeinanderfolgenden Jahren durchgeführten Anbauversuchs an.

---

<sup>1</sup> Vergleiche Kapitel 3.1 der Prüfungsrichtlinien (Dokument TGP/7: Anlage 1: TG-Mustervorlage)

1.2.2.4 Wenn Anbauversuche mit Feld- oder Gewächshauspflanzen in aufeinanderfolgenden Jahren angepflanzt/ausgesät werden, werden sie als unabhängige Wachstumsperioden angesehen.

1.2.2.5 Wenn die beiden Anbauversuche am selben Prüfungsort und im selben Jahr stattfinden, kann ein geeigneter Zeitraum zwischen den Anpflanzungen zwei unabhängige Wachstumsperioden darstellen. Bei Anbauversuchen in Gewächshäusern oder sonstigen stark kontrollierten Umgebungen werden zwei Wachstumsperioden als unabhängige Wachstumsperioden angesehen, sofern der Zeitraum zwischen den beiden Aussaaten nicht „zu kurz“ ist.

1.2.2.6 Wenn zwei Wachstumsperioden im selben Jahr und zur selben Zeit durchgeführt werden, kann ein geeigneter Abstand oder ein angemessener Unterschied bei den Anbaubedingungen zwischen zwei Prüfungsorten die Anforderung der Unabhängigkeit erfüllen.

1.2.2.7 Das Grundprinzip des Einsatzes unabhängiger Wachstumsperioden ist, daß wenn sich der erfaßte Unterschied bei einem Merkmal aus einem genotypischen Unterschied zwischen den Sorten ergibt, dieser Unterschied erfaßt werden sollte, wenn die Sorten erneut in einer ähnlichen Umgebung, jedoch in einer unabhängigen Wachstumsperiode verglichen werden.

## **1.3 Prüfungsort<sup>2</sup>**

### **1.3.1 Zweck**

1.3.1.1 Dokument TGP/7 „Erstellung von Prüfungsrichtlinien“ (vergleiche Anlage I, TG-Mustervorlage, Abschnitt 3.2) stellt klar: „Die Prüfungen werden in der Regel an einem Ort durchgeführt.“ Zu folgenden Zwecken kann es jedoch beispielsweise als angebracht angesehen werden, die Prüfungen an mehr als einem Prüfungsort durchzuführen:

*a) Verkürzung der gesamten Prüfungsperiode*

1.3.1.2 Mehr als ein Prüfungsort kann routinemäßig genutzt werden, beispielsweise als Mittel zur Erzielung von mehr als einer unabhängigen Wachstumsperiode im selben Jahr, wie in 1.2.2.6 dargelegt. Dies könnte die Gesamtdauer der Prüfungsperiode verkürzen und eine raschere Entscheidung erleichtern.

*b) Reserveprüfung*

1.3.1.3 Die Behörden können einen primären Prüfungsort bestimmen, jedoch eine zusätzliche Reserveprüfung an einem anderen Prüfungsort durchführen. In der Regel werden lediglich die Daten aus dem primären Prüfungsort verwendet, doch wenn dieser Prüfungsort ausfällt, wäre die Reserveprüfung verfügbar, um den Verlust der Ergebnisse eines Jahres zu verhindern, sofern es keine signifikante Interaktion zwischen Sorten und Prüfungsort gibt.

---

<sup>2</sup> Vergleiche Kapitel 3.2 der Prüfungsrichtlinien (Dokument TGP/7: Anlage 1: TG-Mustervorlage)

c) *Verschiedene agroklimatische Bedingungen*

1.3.1.4 Verschiedene Sortentypen können verschiedene agroklimatische Anbaubedingungen erfordern. In diesen Fällen müßte der Züchter den Typ der Kandidatensorte angeben, um die Verteilung der Sorte auf den geeigneten Prüfungsort zu ermöglichen. Abschnitt 1.3.2.2 „Zusätzliche Prüfungen“ behandelt die Situation, in der eine Sorte in einer bestimmten Umgebung angebaut werden muß, damit bestimmte Merkmale geprüft werden können, z. B. die Winterfestigkeit. In diesen Fällen wird jedoch jede Sorte an einem Prüfungsort geprüft.

1.3.2 Verwendung von Informationen aus mehreren Prüfungsorten

1.3.2.1 Wenn mehr als ein Prüfungsort genutzt wird, ist es wichtig, Entscheidungsregeln bezüglich der Verwendung von Daten aus verschiedenen Prüfungsorten für die DUS-Prüfung und die Erstellung von Sortenbeschreibungen festzulegen. Zu den Möglichkeiten gehören:

a) *Zusätzliche Prüfungen*

1.3.2.2 Dokument TGP/7 „Erstellung von Prüfungsrichtlinien“ erläutert, daß nebst dem hauptsächlichen Anbauversuch zusätzliche Prüfungen für die Prüfung maßgebender Merkmale durchgeführt werden können (vergleiche Dokument TGP/7: Anlage 1: TG Mustervorlage Abschnitt 3.6). Zusätzliche Prüfungen können beispielsweise durchgeführt werden, um bestimmte Merkmale zu prüfen, z. B. Prüfungen auf Krankheitsresistenz im Gewächshaus, Labortest für chemische Bestandteile usw. In diesen Fällen können die Daten für bestimmte Merkmale an einem anderen Prüfungsort als im Hauptanbauversuch beschafft werden.

b) *DUS aufgrund von Daten, die für dieselben Merkmale an verschiedenen Prüfungsorten erfaßt wurden*

1.3.2.3 Zur Verkürzung der gesamten Prüfungsperiode, wenn zwei unabhängige Wachstumsperioden empfohlen werden (vergleiche Abschnitt 1.3.1 a)), könnte ein zweiter Prüfungsort genutzt werden, um die Stabilität eines am ersten Prüfungsort erfaßten Unterschieds zu überprüfen. Diese Fälle würden in der Regel gelten, wenn die Prüfung der Unterscheidbarkeit auf Noten beruht (vergleiche Dokument TGP/9 Abschnitte 5.2.1.1 b) und 5.2.3) und die Prüfung der Unterscheidbarkeit als auf dem ersten Prüfungsort basierend angesehen werden könnte.

1.3.2.4 Wenn die Prüfung der Unterscheidbarkeit auf der statistischen Analyse der Daten aus dem Anbauversuch beruht, die in zwei oder mehreren unabhängigen Wachstumsperioden erzielt wurden (vergleiche Dokument TGP/9 Abschnitte 5.2.1.1 c) und 5.2.4), könnte es als erwünscht angesehen werden, die Daten aus verschiedenen Prüfungsorten anstatt aus verschiedenen Jahren zu kombinieren, um die gesamte Prüfungsperiode zu verkürzen oder die Daten aus einer Reserveprüfung verwenden zu können. Die Eignung eines solchen Ansatzes würde von den Besonderheiten der betreffenden Pflanzenart abhängen (vergleiche Abschnitt 1.2). Insbesondere müßte sorgfältig geprüft werden, ob die erforderlichen Annahmen erfüllt würden. Diesbezüglich ist anzumerken, daß das COYD-Kriterium an Daten für verschiedene Jahre, nicht an Daten aus verschiedenen Prüfungsorten geprüft wurde.

## **1.4 Bedingungen für die Durchführung der Prüfung**<sup>3</sup>

Dokument TGP/7 „Erstellung von Prüfungsrichtlinien“ erläutert: „Die Prüfungen sollten unter Bedingungen durchgeführt werden, die eine für die Ausprägung der maßgebenden Merkmale der Sorte und für die Durchführung der Prüfung zufriedenstellende Pflanzenentwicklung sicherstellen.“ Spezifische Anleitung wird gegebenenfalls in den entsprechenden Prüfungsrichtlinien erteilt.

## **1.5 Prüfungsanlage**<sup>4</sup>

### 1.5.1 Einleitung

Die DUS-Prüfung beruht im allgemeinen hauptsächlich auf einem Anbauversuch. Es kann zusätzliche Anbauversuche für die Prüfung bestimmter Merkmale oder bestimmter Aspekte von DUS geben, z. B. Ährenreihen für die Prüfung der Homogenität oder zusätzliche Feldprüfungen mit Pflanzen in verschiedenen Entwicklungsstadien, wie junge und ausgewachsene Bäume. Zudem kann es Merkmale geben, die durch zusätzliche Prüfungen geprüft werden müssen, z. B. die Krankheitsresistenz. Die in den folgenden Abschnitten enthaltenen Erläuterungen sollen Anleitung zu den auf Anbauversuche angewandten Grundsätzen geben.

### 1.5.2 Anzahl Pflanzen im Versuch

Die Anzahl Pflanzen im Versuch wird durch mehrere Faktoren beeinflusst, wie der genetischen Struktur der Sorte, der Vermehrungsweise der Art, agronomischer Besonderheiten und der „Durchführbarkeit“ des Versuchs. Die ausschlaggebendsten Kriterien zur Bestimmung der Anzahl Pflanzen sind die Variabilität innerhalb und zwischen Sorten und das Verfahren zur Prüfung der Unterscheidbarkeit und der Homogenität.

### 1.5.3 Prüfungsanlage

#### 1.5.3.1 Einleitung

1.5.3.1.1 Die Art der Prüfungsanlage wird durch das für die Prüfung der Unterscheidbarkeit, der Homogenität und der Beständigkeit angewandte Verfahren bestimmt. Das für die Prüfung der Unterscheidbarkeit angewandte Verfahren ist in Dokument TGP/9 „Prüfung der Unterscheidbarkeit“, Abschnitt 5.2.1, erläutert:

#### „5.2.1 Einleitung

5.2.1.1 Die Verfahren zur Prüfung der Unterscheidbarkeit aufgrund der Anbauprüfung lassen sich wie folgt zusammenfassen:

- a) visueller Seite-an-Seite-Vergleich in der Anbauprüfung (vergleiche Abschnitt 5.2.2);
- b) Prüfung mittels Noten / einmalige Erfassungen von Sorten („Noten“): Die Prüfung der Unterscheidbarkeit beruht auf der erfaßten Ausprägungsstufe der Merkmale der Sorte (vergleiche Abschnitt 5.2.3);

<sup>3</sup> Vergleiche Kapitel 3.3 der Prüfungsrichtlinien (Dokument TGP/7: Anlage 1: TG-Mustervorlage)

<sup>4</sup> Vergleiche Kapitel 3.4 der Prüfungsrichtlinien (Dokument TGP/7: Anlage 1: TG-Mustervorlage)

c) Statistische Analyse der Daten aus der Anbauprüfung: Die Prüfung der Unterscheidbarkeit beruht auf einer statistischen Analyse der aus der Anbauprüfung gewonnenen Daten. Dieses Verfahren setzt voraus, daß es für ein Merkmal eine ausreichende Anzahl Erfassungen für eine Sorte gibt (vergleiche Abschnitt 5.2.4).

5.2.1.2 Die Wahl des Verfahrens oder einer Kombination von Verfahren für die Prüfung der Unterscheidbarkeit wird von der Beobachtungsmethode und der Art der Erfassung (VG, MG, VS oder MS) abhängen, die durch die Besonderheiten der Vermehrung der Sorte und den Ausprägungstyp des Merkmals bestimmt werden. Die üblichen Situationen sind in der Tabelle in Abschnitt 4.5 zusammengefaßt. [...]“

1.5.3.1.2 Die für die Prüfung der Homogenität zu befolgenden Verfahren sind in Dokument TGP/10 „Prüfung der Homogenität“, Abschnitt 2.5.1, erläutert:

„2.5.1 Der Typ der Variation bei der Ausprägung eines Merkmals innerhalb einer Sorte bestimmt, wie dieses Merkmal für die Bestimmung der Homogenität der Pflanze benutzt wird. Wenn es möglich ist, die Abweicher zu ‚visualisieren‘, wird das Abweicherverfahren für die Prüfung der Homogenität empfohlen. Ansonsten wird das Verfahren der Standardabweichungen befolgt. Somit kann die Homogenität einer Sorte entweder nur aufgrund der Abweicher, nur aufgrund der Standardabweichungen oder aufgrund der Abweicher für einige Merkmale und der Standardabweichungen für andere Merkmale bestimmt werden. Diese Situationen werden in Abschnitt 6 weiter untersucht.“

1.5.3.1.3 Dokument TGP/7 „Erstellung von Prüfungsrichtlinien“ ASW 5 – Parzellengestaltung weist folgende Arten der DUS-Prüfung aus:

ASW 5 (TG-Mustervorlage: Kapitel 3.4) – Parzellengestaltung

a) *Einzelparzellen*

„Jede Prüfung sollte so gestaltet werden, daß sie insgesamt mindestens {...} [Pflanzen] / [Bäume] umfaßt.“

b) *Einzelpflanzen und Parzellen in Reihen*

„Jede Prüfung sollte so gestaltet werden, daß sie insgesamt mindestens {...} Einzelpflanzen und {...} Meter Parzellen in Reihen umfaßt.“

c) *Wiederholte Parzellen*

„Jede Prüfung sollte so gestaltet werden, daß sie insgesamt mindestens {...} Pflanzen umfaßt, die auf {...} Wiederholungen aufgeteilt werden sollten.“

Einzelpflanzen und Parzellen in Reihen sind verschiedene Anbauversuche und stellen insbesondere keine wiederholten Parzellen dar (vergleiche Abschnitt 1.5.3.3).

1.5.3.1.4 Anbauversuche ohne Wiederholung (Einzelparzellen) sind geeignet, wenn die Unterscheidbarkeit durch einen visuellen Seite-an-Seite-Vergleich oder durch Noten/einmalige Erfassung von Sorten (vergleiche Dokument TGP/9 Abschnitt 4.3.2.3) geprüft wird und wenn die Homogenität anhand von Abweichern geprüft wird. Übliche Beispiele hierfür sind vegetativ vermehrte Zier- und Obstsorten.

1.5.3.1.5 Wiederholte Parzellen sind geeignet, wenn die Prüfung der Unterscheidbarkeit zumindest für einige Merkmale die Berechnung eines Sortenmittelwerts durch Beobachtung

oder Messung von Gruppen von Pflanzen erfordert (vergleiche Dokument TGP/9 Abschnitt 4.3.2.4). In diesen Fällen wird die Homogenität im allgemeinen anhand von Abweichern geprüft. Übliche Beispiele hierfür sind selbstbefruchtende landwirtschaftliche Pflanzen (z. B. Getreide).

1.5.3.1.6 Wiederholte Parzellen sind geeignet, wenn Erfassungen für eine Anzahl von Einzelpflanzen oder Pflanzenteilen für die statistische Analyse der Daten der Einzelpflanzen für die Prüfung der Unterscheidbarkeit zumindest für einige Merkmale erforderlich sind (in der Regel quantitative Merkmale) (vergleiche Dokument TGP/9 Abschnitt 4.3.3). In diesen Fällen wird die Homogenität für die entsprechenden Merkmale im allgemeinen anhand der Standardabweichung geprüft. Übliche Beispiele hierfür sind fremdbefruchtende Sorten.

1.5.3.1.7 Die nachstehende Tabelle faßt die üblichen Arten der Prüfungsanlage gemäß dem Verfahren für die Prüfung der Unterscheidbarkeit und der Homogenität zusammen:

		HOMOGENITÄT	
		Abweicherverfahren	Standardabweichung
<b>UNTERSCHIEDBARKEIT</b>	Visueller Seite-an-Seite-Vergleich (VG)	Einzelparzellen (vergleiche Abschnitt 1.5.3.2)	
	Noten / einmalige Erfassung von Sorten (VG/MG)	Einzelparzellen (vergleiche Abschnitt 1.5.3.2)	
	Sortenmittelwert Statistische Analyse der Daten für Erfassungen einer Gruppe Pflanzen [Wiederholte Parzellen für Erfassungen von Daten für Gruppen] (MG/MS)	Wiederholte Parzellen	
	Statistische Analyse der Daten von Einzelpflanzen (MS)		Wiederholte Parzellen (vergleiche Abschnitt 1.5.3.3)

MG: einmalige Messung einer Gruppe von Pflanzen oder Pflanzenteilen

MS: Messung einer Anzahl von Einzelpflanzen oder Pflanzenteilen

VG: visuelle Erfassung durch einmalige Beobachtung einer Gruppe von Pflanzen oder Pflanzenteilen

VS: visuelle Erfassung durch Beobachtung einer Anzahl von Einzelpflanzen oder Pflanzenteilen

(vgl. Dokument TGP/9 "Prüfung der Unterscheidbarkeit", Abschnitt 4 "Beobachtung der Merkmale" und TGP/7, Anlage I-TG-Mustervorlage, ASW 7 b).

1.5.3.1.8 Gelegentlich kann es, wie unter den in Dokument TGP/9 Abschnitt 6.4 beschriebenen Umständen angebracht sein, eine randomisierte „Blind“-Prüfung durchzuführen. In diesen Fällen können die bestehenden Parzellen oder Pflanzenteile aus dem Anbauversuch verwendet werden (z. B. die in Dokument TGP/9 Abschnitt 6.4.4 erwähnten ‚randomisiert angeordneten Sortenparzellen‘ und ‚Pflanzenteile von Sorten‘). In anderen Fällen müssen die Pflanzen spezifisch für die randomisierte „Blind“-Prüfung angebaut werden, wie Parzellen mit Pflanzen der beiden Sorten, die zu unterscheiden sind, wobei die Pflanzen in randomisierter, jedoch bekannter Anordnung angebaut werden. In diesem Fall bilden diese Mischparzellen physisch einen Teil der Feldprüfung. Andernfalls kann die randomisierte „Blind“-Prüfung in Form einer Mischung von Töpfen mit den beiden Sorten in einem Gewächshaus erfolgen, was ebenfalls als eine Erweiterung des Anbauversuchs

angesehen wird. Die Anlage dieser randomisierten „Blind“-Prüfung wird in Abschnitt 1.5.3.4 behandelt.

### 1.5.3.2 Einzelparzellen

Diese Prüfungsanlage bedeutet, daß für jede Sorte im Anbauversuch nur eine Parzelle vorhanden ist und daß die Unterscheidbarkeit und die Homogenität in derselben Parzelle geprüft werden.

### 1.5.3.3 Wiederholte Parzellen (statistische Analyse)

#### *1.5.3.3.1 Einleitung*

Wiederholte Parzellen werden benutzt, wenn mehr als eine einmalige Erfassung je Sorte für die Prüfung der Unterscheidbarkeit erforderlich ist. Die Daten aus einer Gruppe von Pflanzen können für die Berechnung eines Sortenmittelwerts verwendet werden, oder die Daten der Einzelpflanzen können für die statistische Analyse verwendet werden.

#### *1.5.3.3.2 Wiederholte Parzellen für die statistische Analyse der Daten von Einzelpflanzen*

1.5.3.3.2.1 Wenn die Prüfung der Unterscheidbarkeit und der Homogenität auf der statistischen Analyse der Daten von Einzelpflanzen beruht, setzt sich der Anbauversuch aus einer Reihe von Parzellen zusammen. Die Parzellen werden im allgemeinen in Wiederholungen gruppiert, so daß jede Wiederholung eine Parzelle jeder Sorte enthält. Die Zuordnung der Sorten an die Parzellen beinhaltet eine Randomisierung (vergleiche Abschnitt 1.5.3.3.3). Beispiele für Prüfungsanlagen, die benutzt werden, wenn eine solche statistische Analyse angewandt wird, sind:

- vollständig randomisierte Anlage und randomisierte vollständige Blockanlage (vergleiche Abschnitt 1.5.3.3.3)
- Randomisierte unvollständige Blockanlage (vergleiche Abschnitt 1.5.3.3.4)
- Anlage für paarweise Vergleiche zwischen bestimmten Sorten (vergleiche Abschnitt 1.5.3.3.5)

1.5.3.3.2.2 Die Unterscheidbarkeit kann für alle Merkmale mittels der statistischen Analyse oder gegebenenfalls für einige Merkmale (insbesondere quantitative Merkmale) durch statistische Analyse und für andere Merkmale (im allgemeinen pseudoqualitative und qualitative Merkmale) durch visuellen Seite-an-Seite-Vergleich oder durch Noten/einmalige Erfassung von Sorten geprüft werden.

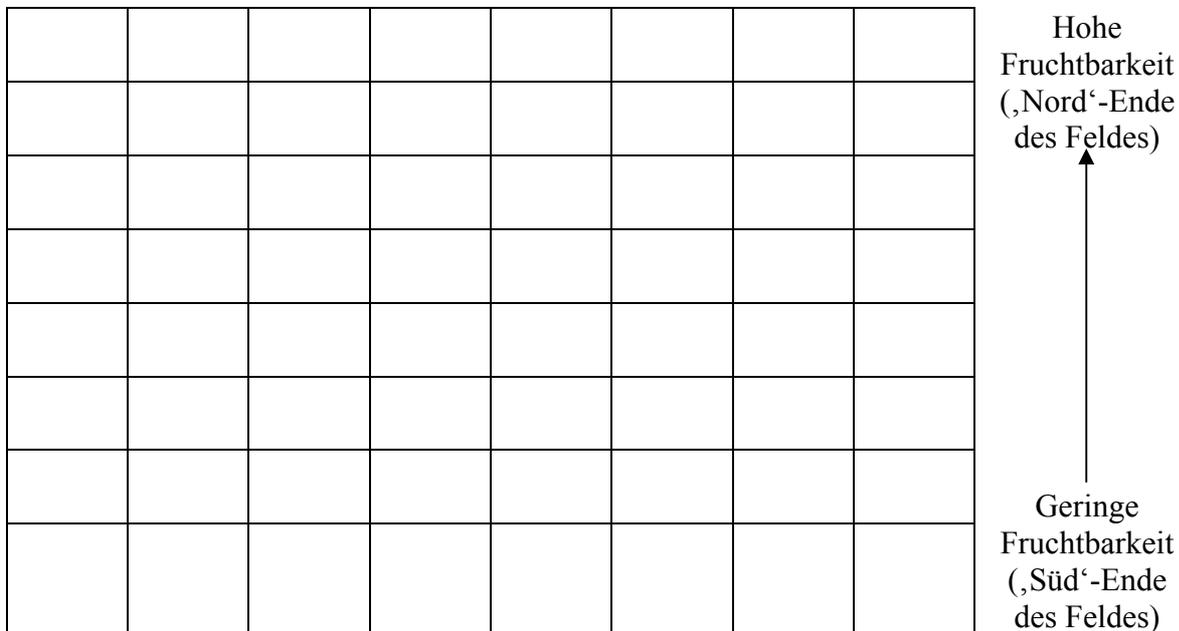
1.5.3.3.2.3 Die Homogenität kann für alle Merkmale mittels der Standardabweichung oder gegebenenfalls für einige Merkmale anhand der Standardabweichung und für andere Merkmale anhand der Abweicher geprüft werden (vergleiche Dokument TGP/10/1, Abschnitt 6.4).

#### *1.5.3.3.3 Randomisierung*

1.5.3.3.3.1 Wenn es im Anbauversuch für jede Sorte wiederholte Parzellen geben soll, ist zu entscheiden, ob die wiederholten Parzellen in Blöcke gruppiert werden sollten und wie die

Parzellen innerhalb eines Block angeordnet werden sollten, d. h. über die Prüfungsanlage. Dies bestimmt, wie die örtliche unerwünschte oder störende Variation kontrolliert wird und daher auch, wie genau die Unterscheidbarkeit und die Homogenität geprüft werden können. Sodann ist die Kenntnis vorhanden, daß sich die Variation aus verschiedenen Quellen ergibt und wie dies die Auswahl der Stichprobengrößen beeinträchtigen kann, was sich wiederum auf die Genauigkeit auswirkt. Die Genauigkeit ist wichtig, weil sie wiederum die Entscheidungsfindung beeinflusst. Wenn die Daten verhältnismäßig ungenau sind und die Entscheidungen auf diesen Daten beruhen, ist die Wahrscheinlichkeit beträchtlich, daß unangemessene oder falsche Entscheidungen getroffen werden. Dies wird nachstehend behandelt.

1.5.3.3.2 Bei der Prüfungsanlage ist es wichtig, eine Fläche zu wählen, die möglichst homogen ist, um die Variation zwischen den Parzellen derselben Sorte, d. h. die Zufallsvariation, zu verringern. Nehmen wir ein Feld an, von dem bekannt ist, daß die größte Variabilität in der ‚Nord-Süd‘-Richtung auftritt, beispielsweise wie in der nachstehenden Abbildung:



1.5.3.3.3 Nehmen wir ein Beispiel, in dem vier Sorten in einem Versuch auf diesem Feld miteinander verglichen werden müssen, bei der jede der Sorten 4 verschiedenen Parzellen zugeordnet wird. Es ist wichtig, die Sorten nach dem Zufallsprinzip auf die Parzellen zu verteilen. Wenn die Sorten systematisch angeordnet werden, hätten nicht alle Sorten zwangsläufig die gleichen Bedingungen (vergleiche nachstehende Abbildung).

Sorte A	Sorte A	Sorte A	Sorte A	Sorte B	Sorte B	Sorte B	Sorte B	Reihe mit höherer Fruchtbarkeit
Sorte C	Sorte C	Sorte C	Sorte C	Sorte D	Sorte D	Sorte D	Sorte D	Reihe mit geringerer Fruchtbarkeit

Wenn die Fruchtbarkeit des Bodens vom Nordende bis zum Süden des Feldes abnimmt, wachsen die Pflanzen der Sorten A und B auf fruchtbareren Parzellen als die übrigen Sorten. Der Vergleich der Sorten wird durch die unterschiedliche Fruchtbarkeit der Parzellen

beeinflußt. Die Unterschiede zwischen den Sorten werden als mit den Fruchtbarkeitsunterschieden konfundiert bezeichnet.

1.5.3.3.3.4 Um systematische Fehler zu vermeiden, ist es ratsam, die Sorten nach dem Zufallsprinzip über das Gelände zu verteilen. Eine vollständige Randomisierung der vier Sorten über die sechzehn Parzellen könnte zu folgender Anlage geführt haben:

Sorte C	Sorte A	Sorte A	Sorte B	Sorte C	Sorte D	Sorte B	Sorte C	Reihe mit höherer Fruchtbarkeit
Sorte C	Sorte A	Sorte D	Sorte A	Sorte D	Sorte B	Sorte D	Sorte B	

1.5.3.3.3.5 Bei Betrachtung der Anlage ist jedoch festzustellen, daß die Sorte C in der oberen Reihe (mit hoher Fruchtbarkeit) dreimal und in der zweite Reihe (mit geringerer Fruchtbarkeit) nur einmal vorkommt. Für die Sorte D ist die Situation umgekehrt. Da wir wissen, daß ein Fruchtbarkeitsgefälle vorhanden ist, ist dies noch immer keine gute Anlage, jedoch besser als die erste systematische Anlage.

1.5.3.3.3.6 Wenn bekannt ist, daß es in den obigen Absätzen gewisse systematische Variationsquellen wie das Fruchtbarkeitsgefälle gibt, kann diese Information berücksichtigt werden, indem sogenannte Blöcke gebildet werden. Die Blöcke sollten so gebildet werden, daß die Variation in jedem der Blöcke minimiert wird. Angesichts des angenommenen Gefälles können entweder zwei Blöcke gewählt werden, die aus je einer Reihe bestehen, oder aber vier Blöcke – zwei Blöcke in jeder Reihe mit je vier Parzellen. Bei größeren Anbauversuchen (mehr Parzellen) ist die letztere Lösung zumeist die beste, da es auch innerhalb der Reihen eine gewisse Variation geben wird, obwohl das größte Gefälle zwischen zwei Reihen besteht.

Block I				Block II				Reihe mit höherer Fruchtbarkeit
Sorte A	Sorte C	Sorte D	Sorte B	Sorte A	Sorte C	Sorte D	Sorte B	
Sorte B	Sorte C	Sorte A	Sorte D	Sorte C	Sorte A	Sorte D	Sorte B	Reihe mit geringerer Fruchtbarkeit
Block III				Block IV				

Eine andere Möglichkeit zur Reduzierung des Effekts eines Gefälles zwischen den Spalten ist die Benutzung von Parzellen, die halb so breit sind, sich jedoch über zwei Reihen erstrecken, d. h. indem lange und schmale Parzellen benutzt werden:

Block I				Block II				Block III				Block IV			
Sorte A	Sorte C	Sorte D	Sorte B	Sorte A	Sorte C	Sorte D	Sorte B	Sorte B	Sorte C	Sorte A	Sorte D	Sorte C	Sorte A	Sorte D	Sorte B

Bei beiden obigen Anlagen beeinträchtigt die ‚Nord-Süd‘-Variabilität die Vergleiche zwischen Sorten nicht.

1.5.3.3.3.7 In einer randomisierten vollständigen Blockanlage ist die Zahl der Parzellen gleich wie die Anzahl Sorten. Alle Sorten sind in jedem Block einmal vorhanden, und die Reihenfolge der Sorten innerhalb jedes Blocks ist randomisiert. Der Vorteil einer randomisierten vollständigen Blockanlage ist, daß die Standardabweichung zwischen den Parzellen (Sorten) – eine Meßgröße der Zufallsvariation – keine Variation infolge der Unterschiede zwischen Blöcken enthält. Der Hauptgrund für die Zuordnung nach dem Zufallsprinzip ist, daß sie sicherstellt, daß die Ergebnisse nicht verzerrt werden und somit die Sorten darstellt, die verglichen werden. Mit anderen Worten reflektieren die Sortenmittelwerte im Durchschnitt die wirklichen Sorteneffekte und werden nicht dadurch aufgebläht oder reduziert, daß sie an sich schon besseren oder schlechteren Parzellen zugeordnet wurden. Eine interessante Besonderheit der Randomisierung ist, daß sie bewirkt, daß sich die Erfassungen aus Einzelparzellen als unabhängige Erfassungen ‚verhalten‘ (obwohl dies vielleicht nicht der Fall ist). In der Regel sind mit der Blockbildung keine zusätzlichen Kosten verbunden. Deshalb wird empfohlen, die Parzellen in Blöcken anzuordnen.

1.5.3.3.3.8 Die Blockbildung wird hier aufgrund der Unterschiede bei der Fruchtbarkeit eingeführt. Mehrere andere systematische Variationsquellen hätten als Grundlage für die Blockbildung dienen können. Obwohl nicht immer klar ist, wie heterogen das Feld ist, und deshalb nicht bekannt ist, wie die Blöcke anzuordnen sind, ist es in der Regel eine gute Idee, Blöcke aus anderen Gründen zu bilden. Wenn mehrere Sämaschinen, verschiedene Erfasser, verschiedene Beobachtungstage vorhanden sind, sind solche Effekte, wenn sie den Parzellen nach dem Zufallsprinzip zugeordnet werden, in der Rest-Standardabweichung enthalten. Diese Effekte können jedoch aus der Rest-Standardabweichung entfernt werden, wenn alle Parzellen innerhalb jedes Blocks dieselbe Sämaschine, denselben Erfasser, denselben Beobachtungstag usw. haben.

1.5.3.3.3.9 Die Verwaltung kann die Wahl der Form der Parzellen beeinflussen. Bei einigen Pflanzen ist es vielleicht leichter, lange, schmale Parzellen als quadratische Parzellen zu handhaben. Lange, schmale Parzellen gelten in der Regel als stärker von Sorten in angrenzenden Parzellen beeinflusst als quadratische Parzellen. Die Parzellengröße sollte so gewählt werden, daß die erforderliche Anzahl Pflanzen für die Stichprobenerhebung vorhanden ist. Für einige Pflanzen kann es notwendig sein, auch Schutzpflanzen (Umfassungstreifen) zu haben, um starke Konkurrenzeffekte zu vermeiden. Übergroße Parzellen benötigen jedoch mehr Land und erhöhen häufig die Zufallsvariabilität zwischen Parzellen. Der gemeinsame Anbau physisch ähnlicher Sorten, beispielsweise Sorten von ähnlicher Höhe, kann die Konkurrenz zwischen benachbarten Parzellen ebenfalls reduzieren. Wenn nichts über die Fruchtbarkeit des Geländes bekannt ist, sind Anlagen mit kompakten Blöcken (d. h. nahezu quadratischen Blöcken) häufig am geeignetsten, weil sie um so unterschiedlicher sind, je größer der Abstand zwischen zwei Parzellen ist. Bei beiden obigen Anlagen können die Blöcke wie angegeben angeordnet werden, oder sie könnten ‚übereinander‘ angeordnet werden (vergleiche nachstehende Abbildung). Dadurch wird die Variabilität zwischen Parzellen in der Regel nicht nennenswert verändert – es sei denn, daß eine der Anlagen den Pflanzensachverständigen dazu zwingt, einen heterogeneren Boden zu benutzen.

Sorte A	Sorte C	Sorte D	Sorte B	Block I	Reihe mit höherer Fruchtbarkeit
Sorte A	Sorte C	Sorte D	Sorte B	Block II	
Sorte B	Sorte C	Sorte A	Sorte D	Block III	Reihe mit geringerer Fruchtbarkeit
Sorte C	Sorte A	Sorte D	Sorte B	Block IV	

#### 1.5.3.3.4 Randomisierte unvollständige Blockanlagen

1.5.3.3.4.1 Wenn die Anzahl Sorten sehr groß wird (>20-40), ist es vielleicht unmöglich, vollständige Blöcke anzulegen, die hinreichend homogen sind. In diesem Fall könnte es von Vorteil sein, kleinere Blöcke zu bilden, von denen jeder lediglich einen Bruchteil der Gesamtzahl der Sorten enthält. Diese Anlagen werden als unvollständige Blockanlagen bezeichnet. In der Literatur sind mehrere Arten unvollständiger Blockanlagen zu finden, beispielsweise ausgewogene unvollständige Blockanlagen und teilweise ausgewogene unvollständige Blockanlagen, wie Gitteranlagen und Reihen- und Säulenanlagen. Eine der bekanntesten Arten für Sortenversuche ist eine Gitteranlage. Die verallgemeinerten Gitteranlagen (auch als  $\alpha$ -Anlagen bezeichnet) sind äußerst flexibel und können für jede beliebige Anzahl Sorten und für eine große Spannweite von Blockgrößen und eine hohe Anzahl Wiederholungen angelegt werden. Eine der Besonderheiten der verallgemeinerten Gitteranlagen ist, daß die unvollständigen Blöcke eine ganze Wiederholung bilden. Das bedeutet, daß diese Anlagen mindestens so gut wie randomisierte vollständige Blockanlagen sind, da die Analyse entweder anhand eines Gittermodells oder eines randomisierten unvollständigen Blockmodells durchgeführt werden kann. Wenn die Bedingungen erfüllt sind, sollte das Gittermodell vorgezogen werden. Die Bestimmung der optimalen Größe der Unterblöcke hängt von verschiedenen Faktoren wie der Variabilität des Bodens und der unterschiedlichen Anfälligkeiten der Merkmale für diese Variabilität ab. Wenn jedoch keine Informationen vorliegen, beispielsweise aus dem ersten Anbauversuch, könnte die anwendbare Zahl der Unterblöcke als Ganzzahl berechnet werden, die der Quadratwurzel der Anzahl Sorten nahe ist; 100 Sorten würden z. B. 10 Unterblöcke erfordern.

1.5.3.3.4.2 Unvollständige Blöcke müssen so angelegt werden, daß es möglich ist, alle Sorten effizient zu vergleichen. Ein Beispiel für eine  $\alpha$ -Anlage ist in der nachstehenden Abbildung gezeigt:

Block I	Unterblock I	Sorte F	Sorte E	Sorte O	Sorte S
	Unterblock II	Sorte M	Sorte H	Sorte J	Sorte T
	Unterblock III	Sorte B	Sorte C	Sorte D	Sorte G
	Unterblock IV	Sorte L	Sorte A	Sorte R	Sorte N
	Unterblock V	Sorte Q	Sorte K	Sorte P	Sorte I

Block II	Unterblock I	Sorte D	Sorte P	Sorte F	Sorte A
	Unterblock II	Sorte R	Sorte E	Sorte J	Sorte B
	Unterblock III	Sorte N	Sorte G	Sorte Q	Sorte H
	Unterblock IV	Sorte K	Sorte S	Sorte M	Sorte C
	Unterblock V	Sorte O	Sorte I	Sorte T	Sorte L

Block III	Unterblock I	Sorte D	Sorte T	Sorte E	Sorte Q
	Unterblock II	Sorte B	Sorte M	Sorte A	Sorte I
	Unterblock III	Sorte C	Sorte F	Sorte L	Sorte H
	Unterblock IV	Sorte R	Sorte G	Sorte K	Sorte O
	Unterblock V	Sorte P	Sorte J	Sorte N	Sorte S

Im obigen Beispiel sollen 20 Sorten in einem Anbauversuch mit drei Wiederholungen angebaut werden. In der Anlage bilden die 5 Unterblöcke jedes Blocks eine vollständige Wiederholung. Somit enthält jede Wiederholung alle Sorten, während ein beliebiges Sortenpaar im selben Unterblock entweder einmal oder überhaupt nicht vorkommt. Anmerkung: In der Literatur werden die Blöcke und Unterblöcke mitunter als Superblöcke und Blöcke bezeichnet.

1.5.3.3.4.3 Die unvollständige Blockanlage ist für Anbauversuche am geeignetsten, bei denen keine Gruppierungsmerkmale verfügbar sind. Wenn Gruppierungsmerkmale verfügbar sind, kann eine gewisse Änderung für Anbauversuche mit zahlreichen Sorten vorteilhaft sein, wie die Verwendung von Gruppierungsmerkmalen zur Bildung getrennter Anbauversuche anstatt eines einzigen Anbauversuchs; vergleiche Dokument TGP/9/1, Abschnitt 2.3, Gruppierung von Sorten aufgrund von Merkmalen.

#### 1.5.3.3.5 Anlage für Paarvergleiche zwischen bestimmten Sorten

1.5.3.3.5.1 Wenn ein genauer Vergleich zwischen einem Sortenpaar mittels der statistischen Analyse benötigt wird, kann es angemessen sein, sie in benachbarten Parzellen anzubauen. Eine ähnliche Theorie wie diejenige, die bei Spaltenparzellenanlagen angewandt wird, kann für die Einrichtung einer Anlage angewandt werden, bei der die Vergleiche zwischen bestimmten Sortenpaaren optimiert werden sollen. Bei der Einrichtung der Anlage werden die Sortenpaare als gesamter Parzellenfaktor behandelt, und der Vergleich zwischen Sorten

innerhalb jedes Paares ist der Unterparzellenfaktor. Da jede Parzelle lediglich aus zwei Unterparzellen besteht, sind die Vergleiche innerhalb von Paaren (viel) genauer, als wenn eine randomisierte Blockanlage angewandt würde.

1.5.3.3.5.2 Wenn beispielsweise vier Sortenpaare (A-B, C-D, E-F und G-H) sehr genau verglichen werden müssen, kann dies anhand der folgenden Anlage von 12 ganzen Parzellen mit je 2 Unterparzellen erfolgen:

Paar 1 Sorte A	Paar 3 Sorte E	Paar 4 Sorte H
Paar 1 Sorte B	Paar 3 Sorte F	Paar 4 Sorte G
Paar 3 Sorte F	Paar 2 Sorte D	Paar 1 Sorte A
Paar 3 Sorte E	Paar 2 Sorte C	Paar 1 Sorte B
Paar 4 Sorte G	Paar 1 Sorte B	Paar 2 Sorte C
Paar 4 Sorte H	Paar 1 Sorte A	Paar 2 Sorte D
Paar 2 Sorte D	Paar 4 Sorte H	Paar 3 Sorte E
Paar 2 Sorte C	Paar 4 Sorte G	Paar 3 Sorte F

In dieser Anlage stellt jede Spalte eine Wiederholung dar. Jede davon ist sodann in vier unvollständige Blöcke (ganze Parzellen) aufgeteilt, die je aus zwei Unterparzellen bestehen. Die vier Sortenpaare werden nach dem Zufallsprinzip innerhalb jeder Wiederholung angeordnet, und die Anordnung der Sorten wird innerhalb jedes unvollständigen Blocks randomisiert. Der Vergleich zwischen Sorten desselben Paares wird auf Kosten der Genauigkeit des Vergleiches zwischen Sorten eines verschiedenen Paares genauer gemacht.

#### 1.5.3.3.6 Weitere statistische Aspekte der Prüfungsanlage

##### 1.5.3.3.6.1 Einleitung

1.5.3.3.6.1.1 Dieser Abschnitt beschreibt eine Reihe von Begriffen, die beim Anlegen der Anbauversuche von Belang sind, bei denen die Unterscheidbarkeit und/oder Homogenität durch statistische Analyse der Daten aus dem Anbauversuch geprüft werden sollen.

##### 1.5.3.3.6.2 Die Hypothesenprüfung

1.5.3.3.6.2.1 Wenn die statistische Analyse der Daten aus dem Anbauversuch für die Prüfung der Unterscheidbarkeit und/oder Homogenität angewandt werden soll, ist es der Zweck des Anbauversuchs, genaue und unverzerrte Durchschnitte der Merkmale für jede Sorte zu erzielen und auch die Variabilität innerhalb der Sorten zu beurteilen, indem die Standardabweichung berechnet wird. Die Prüfung der Unterscheidbarkeit der Sorten erfolgt aufgrund der Merkmalsdurchschnitte. Der Typ der Variation bei der Ausprägung eines Merkmals innerhalb einer Sorte bestimmt, wie dieses Merkmal für die Bestimmung der Homogenität der Pflanze benutzt wird. Wenn es möglich ist, die Abweicher zu ‚visualisieren‘, wird das Abweicherverfahren für die Prüfung der Homogenität empfohlen. Ansonsten wird das Verfahren der Standardabweichungen befolgt.

1.5.3.3.6.2.2 Bei der Bewertung der Unterscheidbarkeit und der Homogenität wird eine Nullhypothese ( $H_0$ ) geprüft und entweder akzeptiert oder zurückgewiesen. Wenn sie zurückgewiesen wird, wird eine alternative Hypothese ( $H_1$ ) akzeptiert. Die Null- und die alternative Hypothese für die Entscheidungen über die Unterscheidbarkeit und die Homogenität sind in der nachstehenden Tabelle wiedergegeben:

	Nullhypothese ( $H_0$ )	Alternative Hypothese ( $H_1$ )
<i>Unterscheidbarkeit</i>	zwei Sorten sind für das Merkmal nicht unterscheidbar	zwei Sorten sind unterscheidbar
<i>Homogenität</i>	eine Sorte ist für das Merkmal homogen	eine Sorte ist nicht homogen

1.5.3.3.6.2.3 Jede Bewertung wird durch Berechnung einer Testkenngröße aus den Erfassungen anhand einer Formel vorgenommen. Wenn der absolute Wert der Testkenngröße größer als ihr gewählter kritischer Wert ist, wird die Nullhypothese ( $H_0$ ) zurückgewiesen, die alternative Hypothese ( $H_1$ ) akzeptiert, und der Test wird als signifikant bezeichnet. Wenn die Testkenngröße nicht größer als ihr gewählter kritischer Wert ist, wird die Nullhypothese ( $H_0$ ) akzeptiert. Die Wahl des kritischen Wertes, mit dem die Testkenngröße verglichen wird, ist nachstehend erläutert.

1.5.3.3.6.2.4 Es ist zu beachten, daß wenn die Nullhypothese ( $H_0$ ) für die Unterscheidbarkeit zurückgewiesen wird, dies zur Schlußfolgerung führt, daß die Kandidatensorte unterscheidbar ist.

1.5.3.3.6.2.5 Wenn andererseits die Nullhypothese ( $H_0$ ) für die Homogenität zurückgewiesen wird, wird die Kandidatensorte als nicht homogen angesehen.

1.5.3.3.6.2.6 Die Testkenngröße beruht auf einer Stichprobe von Pflanzen, die in einer Stichprobe von Anbaubedingungen im Anbauversuch angebaut werden. Wenn der Prozeß zu einem anderen Zeitpunkt wiederholt werden müßte, würde daher ein anderer Wert für die Testkenngröße erzielt. Wegen dieser inhärenten Variabilität besteht die Wahrscheinlichkeit, daß im Vergleich zu der Schlußfolgerung, die erreicht würde, wenn der Anbauversuch unbegrenzt wiederholt werden könnte, eine andere Schlußfolgerung erreicht wird. Solche „statistischen Fehler“ können auf zwei Arten auftreten. Wir untersuchen zunächst die Schlußfolgerungen zur Unterscheidbarkeit:

- Die auf der Testkenngröße, d. h. aus der DUS-Prüfung, beruhenden Schlußfolgerungen lauten, daß zwei Sorten unterscheidbar sind, die nicht unterscheidbar wären, falls der Anbauversuch unbegrenzt wiederholt werden könnte. Dies ist als Fehler vom Typ I bekannt, und sein Risiko wird durch  $\alpha$  gekennzeichnet.
- Die auf der Testkenngröße, d. h. aus der DUS-Prüfung, beruhenden Schlußfolgerungen lauten, daß zwei Sorten nicht unterscheidbar sind, die unterscheidbar wären, falls der Anbauversuch unbegrenzt wiederholt werden könnte. Dies ist als Fehler vom Typ II bekannt, und sein Risiko wird durch  $\beta$  gekennzeichnet.

<b>Schlußfolgerung, wenn der Anbauversuch unbegrenzt wiederholt werden könnte</b>	<b>Auf einer Testkenngröße beruhende Schlußfolgerung</b>	
	die Sorten sind nicht unterscheidbar ( $H_0$ trifft zu)	die Sorten sind unterscheidbar ( $H_1$ trifft zu)
Die Sorten sind unterscheidbar ( $H_1$ trifft zu)	Anderes Ergebnis, Fehler vom Typ II mit der Wahrscheinlichkeit $\beta$	gleiches Ergebnis
Die Sorten sind nicht unterscheidbar ( $H_0$ trifft zu)	gleiches Ergebnis	anderes Ergebnis, Fehler vom Typ I, mit der Wahrscheinlichkeit $\alpha$

1.5.3.3.6.2.7 Gleichermaßen ist es bei der Entscheidung über die Homogenität aufgrund einer Testkenngröße, d. h. aus der DUS-Prüfung, möglich zu entscheiden, daß eine Sorte nicht homogen ist, wenn sie, falls der Anbauversuch unbegrenzt wiederholt werden könnte, homogen wäre, d. h. ein Fehler vom Typ I ( $\alpha$ ). Andernfalls ist die Schlußfolgerung aufgrund einer Testkenngröße, daß eine Sorte homogen ist, wenn sie, falls der Anbauversuch unbegrenzt wiederholt werden könnte, nicht homogen wäre, ein Fehler vom Typ II ( $\beta$ ). Die nachstehende Tabelle zeigt die beiden Typen statistischer Fehler, die bei der Prüfung der Homogenität auftreten können:

<b>Schlußfolgerung, wenn der Anbauversuch unbegrenzt wiederholt werden könnte</b>	<b>Auf einer Testkenngröße beruhende Schlußfolgerung</b>	
	die Sorte ist homogen ( $H_0$ trifft zu)	Sorte ist nicht homogen ( $H_1$ trifft zu)
Die Sorte ist homogen ( $H_0$ trifft zu)	gleiches Ergebnis	anderes Ergebnis, Fehler vom Typ I, mit der Wahrscheinlichkeit $\alpha$
Sorte ist nicht homogen ( $H_1$ trifft zu)	anderes Ergebnis, Fehler vom Typ II, mit der Wahrscheinlichkeit $\beta$	gleiches Ergebnis

1.5.3.3.6.2.8 Das Risiko, einen Fehler vom Typ I zu begehen, kann problemlos kontrolliert werden durch die Wahl von  $\alpha$ , das den kritischen Wert bestimmt, mit dem die Testkenngröße verglichen wird.  $\alpha$  ist auch als Größe des Tests und als Signifikanzniveau des Tests bekannt. Das Risiko eines Fehlers vom Typ II ist schwieriger zu kontrollieren, da es beispielsweise im Fall der Unterscheidbarkeit von der Größe des tatsächlichen Unterschieds zwischen den Sorten, dem gewählten  $\alpha$  und der Genauigkeit des Tests abhängt, die durch die Anzahl Wiederholungen und die inhärente Variabilität der Messungen bestimmt wird. Der Pflanzensachverständige kann das Risiko eines Fehlers vom Typ II reduzieren, indem er die Genauigkeit erhöht, beispielsweise durch Erhöhung der Anzahl Wiederholungen, durch Reduzierung der Zufallsvariabilität mittels der Wahl der Anzahl Pflanzen pro Parzelle (oder Stichprobengröße), durch die Kontrolle der örtlichen unerwünschten oder störenden Variation mittels einer sorgfältigen Wahl der Prüfungsanlage und durch die Verbesserung der Art und Weise, wie die Messungen/Erfassungen vorgenommen werden und so der Fehler der Erfasser reduziert wird.

#### 1.5.3.3.6.3 Bestimmung der optimalen Stichprobengröße

1.5.3.3.6.3.1 Die Genauigkeit eines Tests hängt nicht allein von der Stichprobengröße ab. Die Genauigkeit eines auf Beobachtungen eines Versuchs beruhenden Tests hängt beispielsweise für quantitative Merkmale auch von mindestens drei Variationsquellen ab:

- der Variation zwischen Einzelpflanzen innerhalb einer Parzelle, d. h. der Varianzkomponente „innerhalb der Parzelle“ oder der „Pflanzen“-Varianzkomponente: eine Mischung verschiedener Variationsquellen wie verschiedene Pflanzen, verschiedene Beobachtungszeiten, verschiedene Meßfehler
- der Variation zwischen den Parzellen innerhalb eines Blocks, d. h. der Varianzkomponente „zwischen Parzellen“ oder „Parzellen“-Varianzkomponente

- der durch die Umwelt verursachten Variation, d. h. der Variation bei der Ausprägung der Merkmale von Jahr zu Jahr (oder von Prüfungsort zu Prüfungsort)

1.5.3.3.6.3.2 Zur Schätzung der optimalen Stichprobengröße für ein quantitatives Merkmal ist es notwendig, die Standardabweichungen der obigen Variationsquellen, die erwarteten Unterschiede zwischen den Sorten, die signifikant sein sollten, die Anzahl Sorten und die Anzahl Blöcke im Anbauversuch zu kennen. Zudem müssen die Wahrscheinlichkeiten der Fehler vom Typ I ( $\alpha$ ) und vom Typ II ( $\beta$ ) bestimmt werden. Die Berechnung der optimalen Stichprobengröße für jedes Merkmal ermöglicht eine Bestimmung der optimalen Stichprobengröße für diesen Anbauversuch für alle quantitativen Merkmale. Insbesondere für die Prüfung der Homogenität ist der Fehler vom Typ II mitunter wichtiger als der Fehler vom Typ I. In einzelnen Fällen könnte der Fehler vom Typ II größer als 50 % sein, was unzulässig sein kann.

1.5.3.3.6.3.3 Die Genauigkeit der Sortenmittelwerte in einem Anbauversuch in einem Jahr oder in einer Wachstumsperiode hängt von der Anzahl Wiederholungen, der Anzahl Pflanzen pro Parzelle und der Versuchsanlage ab. Wenn diese Mittelwerte beispielsweise in der Analyse über die Jahre oder über die Wachstumsperioden für COYD verwendet werden, ist ihre Genauigkeit nur indirekt von Nutzen, weil die Standardabweichung in dieser Analyse auf der Interaktion zwischen den Sorten und den Jahren oder Wachstumsperioden beruht. Außerdem ist die Genauigkeit der Mittelwerte je Versuch verhältnismäßig unwichtig, wenn die Unterschiede zwischen den Sorten über die Jahre oder die Wachstumsperioden sehr groß sind.

1.5.3.3.6.3.4 Wenn UPOV-Prüfungsrichtlinien vorliegen, empfehlen sie unter Berücksichtigung der oben erläuterten Faktoren eine geeignete Stichprobengröße für den gesamten Anbauversuch.

#### *1.5.3.3.7 Elemente des Anbauversuchs bei Einsatz der statistischen Analyse*

##### *1.5.3.3.7.1 Einleitung*

1.5.3.3.7.1.1 Bei der Entscheidung über die Prüfungsanlage ist es wichtig, daß die örtliche Variation der Bedingungen berücksichtigt wird. Hierfür sind Entscheidungen über die Parzellengröße, die Form der Parzellen, die Ausrichtung der Parzellen, Trennreihen und Rendreihen sowie Randstreifen notwendig.

1.5.3.3.7.1.2 Für die Prüfung der Unterscheidbarkeit ist die unverzerrte Beobachtung der Merkmale notwendig. In einigen Fällen ist es notwendig, Rendreihen und Randstreifen zur Reduzierung der Verzerrung zu haben, die durch die Beeinflussung zwischen Parzellen, d. h. die Beeinflussung zwischen Pflanzen in verschiedenen Parzellen und sonstige besondere Randeffekte wie Beschattung oder Bodenfeuchtigkeit verursacht wird. Zudem werden häufig Schutzstreifen am Rand des Anbauversuchs benutzt, um die Wahrscheinlichkeit äußerer Einflüsse zu reduzieren, die eine Parzelle zugunsten einer anderen verzerren könnten. Bei der Erfassung der Merkmale an den Pflanzen einer Parzelle ist es üblich, die Rendreihen und Randstreifen der Parzelle auszuschließen.



#### *1.5.3.3.7.3 Zuordnung der Sorten zu den Parzellen*

1.5.3.3.7.3.1 Mehrere Faktoren beeinflussen die Entscheidung über die Zuordnung der Sorten zu den Parzellen, insbesondere das gewählte Verfahren für die Unterscheidbarkeit (vergleiche Abschnitt 1.5.3.1.1) und die Homogenität (vergleiche Abschnitt 1.5.3.1.2).

1.5.3.3.7.3.2 Wenn die Unterscheidbarkeit anhand der statistischen Analyse der Daten aus dem Anbauversuch geprüft wird, muß je nach Prüfungsanlage entweder die Randomisierung oder die Teilrandomisierung angewandt werden, da diese sicherstellt, daß es bei der Zuordnung keine Subjektivität gibt. Die randomisierte Zuordnung stellt sicher, daß sich die Effekte anderer Faktoren, die die Merkmale der Pflanzen beeinflussen, wie die Bodenbedingungen, voraussichtlich im Durchschnitt aufheben, wenn die Sortenmittelwerte verglichen werden.

1.5.3.3.7.3.3 Die Abschnitte 1.5.3.2 und 1.5.3.3.1 bis 1.5.3.3.5 geben weitere Einzelheiten über die verschiedenen Arten der Sortenzuordnung zu den Parzellen und Blöcken an.

#### *1.5.3.3.7.4 Größe, Form und Anordnung*

1.5.3.3.7.4.1 Abschnitt 3 der Prüfungsrichtlinien „Durchführung der Prüfung“ erteilt Informationen über die Dauer der Prüfung, den Prüfungsort, die Gestaltung der Prüfung, die Anzahl der zu prüfenden Pflanzen/Pflanzenteile sowie zusätzliche Prüfungen, die für die Prüfung maßgebender Merkmale durchgeführt werden können. Die Prüfungsrichtlinien können die Art der Erfassung für die Prüfung der Unterscheidbarkeit (einmalige Erfassung für eine Gruppe von Pflanzen oder Pflanzenteilen (G) oder Erfassungen für eine Anzahl individueller Einzelpflanzen oder Pflanzenteile (S)) angeben. Die Homogenität wird jedoch an der gesamten zu prüfenden Stichprobe anhand des Abweicherverfahrens und/oder des Standardabweichungsverfahrens geprüft (vergleiche Dokument TGP/10 Abschnitt 3). Diese bestimmen die Stichprobengröße, d. h. die Anzahl Pflanzen, die beobachtet werden müssen, und daher auch die effektive Mindestgröße der Parzelle. Für die Entscheidung über die tatsächliche Parzellengröße müssen etwaige notwendige Rendreihen und -streifen berücksichtigt werden.

1.5.3.3.7.4.2 Die Parzellengröße und die Parzellenform hängen ebenfalls von den Boden- und sonstigen Bedingungen, der Bewässerungsanlage oder den Sä- und Erntemaschinen ab. Die Form der Parzelle läßt sich als das Verhältnis der Parzellenlänge dividiert durch die Parzellenbreite definieren. Dieses Verhältnis kann wichtig sein, um die (z. B. durch Bodenvariation verursachte) Variation bei den Bedingungen innerhalb des Blocks zu mildern.

1.5.3.3.7.4.3 Quadratische Parzellen weisen die kleinste Gesamtlänge der Ränder (Umfang) auf. Vom theoretischen Standpunkt aus ist die quadratische Form optimal für die Minimierung des Einflusses verschiedener Phänotypen. Die Gruppierung der Sorten kann ebenfalls zur Minimierung dieser Beeinflussung beitragen.

1.5.3.3.7.4.4 Schmale, lange Parzellen werden aus technischer Sicht bevorzugt. Das beste Verhältnis Länge/Breite liegt zwischen 5:1 und 15:1 und hängt von der Parzellengröße und der Anzahl Sorten ab. Je größer die Anzahl Sorten in einem Block ist, desto schmaler sind die Parzellen – jedoch nicht so schmal, daß die Konkurrenz zwischen Parzellen zum Problem wird.

#### *1.5.3.3.7.5 Unabhängigkeit der Parzellen*

1.5.3.3.7.5.1 Wenn die Unterscheidbarkeit und die Homogenität durch statistische Analyse der Daten aus dem Anbauversuch geprüft werden sollen, ist die Unabhängigkeit eine der wichtigsten Anforderungen an die Versuchseinheiten.

1.5.3.3.7.5.2 Die Unabhängigkeit der Parzellen bedeutet, daß die Erfassungen an einer Parzelle durch die Verhältnisse in anderen Parzellen nicht beeinflußt werden. Wenn beispielsweise hohe Sorten neben niedrigen gepflanzt werden, könnte es eine negative Beeinflussung der niedrigen durch die hohen und in die andere Richtung eine positive Beeinflussung geben. In diesem Fall kann, um diese Abhängigkeit zu vermeiden, eine zusätzliche Pflanzenreihe auf beiden Seiten der Parzelle, d. h. Randreihen und -streifen, gepflanzt werden. Eine weitere Möglichkeit zur Minimierung dieser Beeinflussung ist der gemeinsame Anbau physisch ähnlicher Sorten.

#### *1.5.3.3.7.6 Anordnung der Pflanzen innerhalb der Parzelle / Typ der Parzelle für die Beobachtung*

Die UPOV-Prüfungsrichtlinien können für die Prüfung der Unterscheidbarkeit sowie der Homogenität und der Beständigkeit den/die Typ/en von Parzellen für den Anbauversuch (z. B. Einzelpflanzen, Parzelle in Reihen, gedrillte Parzelle usw.) angeben.

#### 1.5.3.4 Randomisierte Blindprüfungen

1.5.3.4.1 Ein Teil des Anbauversuchs kann aus Parzellen bestehen, die spezifisch für randomisierte „Blind“-Prüfungen angebaut werden, wie Parzellen mit Pflanzen der beiden Sorten, die unterschieden werden sollen, wobei die Pflanzen in randomisierter, jedoch bekannter Anordnung angebaut werden, oder andernfalls aus einer Mischung von Töpfen mit den beiden Sorten in einem Gewächshaus. Die beiden Sorten umfassen die Kandidatensorte sowie die Sorte, bei der die Unterscheidbarkeit der Kandidatensorte fraglich ist. Der Grundsatz der randomisierten „Blind“-Prüfung ist, daß einem Beurteiler, mitunter auch einem Züchter, die Pflanzen vorgelegt werden und sie ersucht werden, Pflanze um Pflanze zu erklären, welches die Kandidatensorte und welches die andere Sorte ist.

1.5.3.4.2 Um dies zu ermöglichen, müssen die Pflanzen in randomisierter Anordnung vorgelegt oder angebaut werden, jedoch so, daß der Prüfer weiß, welches welche Sorte ist, der Beurteiler jede Sorte beurteilt und der Prüfer zählt, wie viele Male die verschiedenen Sorten richtig identifiziert werden. Zur Verstärkung der Blindheit der Prüfung wird von jeder der beiden Sorten eine unterschiedliche Anzahl Pflanzen vorgelegt, beispielsweise 51 von der Kandidatensorte und 69 von der anderen Sorte, anstatt 60 von jeder Sorte. Da es in verschiedenen Entwicklungsstadien Unterschiede geben kann, kann der Beurteiler die Pflanzen mehr als einmal beurteilen.

## **1.6 Änderung der Verfahren**

Änderungen der Verfahren für die DUS-Prüfung können sich erheblich auf die Entscheidungen auswirken. Deshalb sollte gewissenhaft dafür gesorgt werden, daß die Entscheidungen, die Verfahren zu ändern, übereinstimmen.

## **2. VALIDIERUNG DER DATEN UND ANNAHMEN**

### **2.1 Einleitung**

Es ist wichtig, daß die Daten richtig, d. h. fehlerfrei sind. Dies ist der Fall ungeachtet dessen, ob die Daten Noten aus der visuellen Erfassung (V) (vergleiche Dokument TGP/9 Abschnitt 4.2.1) oder Messung (M) (vergleiche Dokument TGP/9 Abschnitt 4.2.2) sind und ob sie zu einer einmaligen Erfassung für eine Gruppe von Pflanzen oder Pflanzenteilen (G) (vergleiche Dokument TGP/9 Abschnitt 4.3.2) oder zu Erfassungen für eine Anzahl individueller Einzelpflanzen oder Pflanzenteile (S) (vergleiche Dokument TGP/9 Abschnitt 4.3.3) für die statistische Analyse führen. Der Abschnitt „Validierung der Daten“ beschreibt, wie die Daten validiert oder kontrolliert werden können. Diese vorläufigen Kontrollen können für alle Daten vorgenommen werden, ungeachtet dessen, ob sie in der Folge anhand statistischer Verfahren analysiert werden oder nicht.

### **2.2 Validierung der Daten**

2.2.1 Dieser Abschnitt befaßt sich mit der Validierung der Daten, um sicherzustellen, daß es keine (offensichtlichen) Fehler gibt.

2.2.2 Um Fehler bei der Interpretation der Ergebnisse zu vermeiden, sollten die Daten stets kontrolliert werden, so daß sie logisch übereinstimmen und nicht in Widerspruch zu früheren Informationen über die Spannweiten stehen, die sich für die verschiedenen Merkmale ergeben könnten. Diese Kontrolle kann manuell (in der Regel visuell) oder automatisch erfolgen. Wenn statistische Verfahren angewendet werden, kann die Validierung der Annahmen auch als Kontrolle genutzt werden, daß die Daten fehlerfrei sind (vergleiche Abschnitt 2.3.2.1.1.)

2.2.3 Tabelle 1 zeigt einen Auszug aus einigen Erfassungen für 10 Pflanzen aus einer Parzelle mit Futtererbsen. Für ‚Samen: Form‘ (PQ) werden die Noten visuell auf einer Skala mit den Werten 1 (kugelförmig), 2 (eiförmig), 3 (zylindrisch), 4 (rhomboid), 5 (dreieckig) oder 6 (unregelmäßig) angegeben. Für ‚Samen: schwarze Farbe der Narbe‘ (QL) werden die Noten visuell auf einer Skala mit den Werten 1 (fehlend) oder 9 (vorhanden) angegeben. Für ‚Stengel: Länge‘ (QN) sind die Messungen in cm angegeben, und aus früherer Erfahrung ist bekannt, daß die Länge in den meisten Fällen zwischen 40 und 80 cm liegt. ‚Nebenblatt: Länge‘ wird in mm gemessen und liegt in den meisten Fällen zwischen 50 und 90 mm. Die Tabelle zeigt drei Arten von Fehlern, die bei manuellen Erfassungen gelegentlich auftreten: Für Pflanze 4, ‚Samen: Form‘ gehört der erfaßte Wert, 7, nicht zu den zulässigen Noten und muß daher auf einen Fehler zurückzuführen sein. Er könnte durch falsche Lektüre eines handgeschriebenen „1“ entstanden sein. Eine ähnliche Situation ist für Pflanze 8 für das Merkmal ‚Samen: schwarze Farbe der Narbe‘ festzustellen, wo die Note 8 nicht zulässig ist und ein Fehler sein muß. ‚Stengel: Länge‘ der Pflanze 6 ist außerhalb der erwarteten Spannweite und könnte durch eine Änderung der Reihenfolge der Zahlen verursacht worden sein, so daß 96 anstatt 69 eingegeben wurde. ‚Nebenblatt: Länge‘ von 668 mm ist eindeutig falsch. Dies könnte durch versehentliche Wiederholung der Zahl 6 verursacht worden sein. In allen Fällen muß eine sorgfältige Prüfung durchgeführt werden, um festzustellen, welches die richtigen Werte sein sollten.

Tabelle 1 Auszug aus dem Erfassungsblatt für Futtererbse

Pflanze Nr.	Samen: Form (UPOV 1)  (PQ)	Samen: schwarze Farbe der Narbe (UPOV 6) (QL)	Stengel: Länge (UPOV 12)  (QN)	Nebenblatt: Länge (UPOV 31)  (QN)
1	1	1	43	80
2	2	1	53	79
3	1	1	50	72
4	7	1	43	668
5	2	9	69	72
6	1	1	96	72
7	1	1	51	70
8	2	8	64	63
9	1	1	44	62
10	2	1	49	62

2.2.4 Graphische Darstellungen oder Diagramme der Merkmale können die Validierung der Daten unterstützen. Die Prüfung der Häufigkeitsverteilungen der Merkmale kann beispielsweise kleine Gruppen abweichender Erfassungen ermitteln. Zudem kann im Fall quantitativer Merkmale die Prüfung von Streudiagrammen von Merkmalspaaren, die aller Wahrscheinlichkeit nach stark verbunden sind, abweichende Erfassungen äußerst effizient aufdecken.

2.2.5 Weitere Arten graphischer Darstellungen können ebenfalls für die Validierung der Datenqualität angewandt werden. Ein sogenanntes Box-Plot ist ein effizientes Mittel, einen Überblick über quantitative Daten zu gewinnen. In einem Box-Plot wird ein Kasten für jede Gruppe (Parzelle oder Sorte) gezeichnet. In diesem Fall werden Daten von ‚Blatt: Länge‘ (in mm) aus einem in 3 Blöcken von 26 Parzellen mit 20 Pflanzen pro Parzelle angelegten Versuch verwendet. In jedem Block wurden 26 verschiedene Sorten von Raps jeder Parzelle nach dem Zufallsprinzip zugeordnet. In Abbildung 1 werden alle 60 ‚Blatt: Länge‘ jeder der 26 Sorten zusammengenommen. (Wenn es große Blockunterschiede gibt, kann ein besseres Box-Plot erstellt werden, indem die Unterschiede in bezug auf den Parzellenmittelwert verwendet werden). Der Kasten zeigt die Spannweite für den größten Teil der individuellen Erfassungen (in der Regel 75 %). Eine horizontale Linie durch den Kasten und ein Symbol geben den Median bzw. den Mittelwert an. An jedem Ende des Kastens werden vertikale Linien gezogen, um die Spannweite möglicher Erfassungen außerhalb des Kastens, jedoch innerhalb eines angemessenen Abstandes, anzugeben (in der Regel 1,5 Mal die Höhe des Kastens). Schließlich werden extreme Erfassungen einzeln gezeigt. In Abbildung 1 ist festzustellen, daß eine Beobachtung der Sorte 13 eindeutig weit größer als die übrigen Erfassungen dieser Sorte ist. Zudem ist festzustellen, daß die Sorte 16 große Blattlängen aufweist und daß etwa 4 Erfassungen relativ weit vom Mittelwert entfernt sind. Aus der Abbildung sind u. a. die Variabilität und die Symmetrie der Verteilung zu ersehen. Somit ist festzustellen, daß die Variabilität der Sorte 15 verhältnismäßig groß ist und daß die Verteilung für diese Sorte etwas schief ist (da der Mittelwert und der Median relativ weit voneinander entfernt sind).

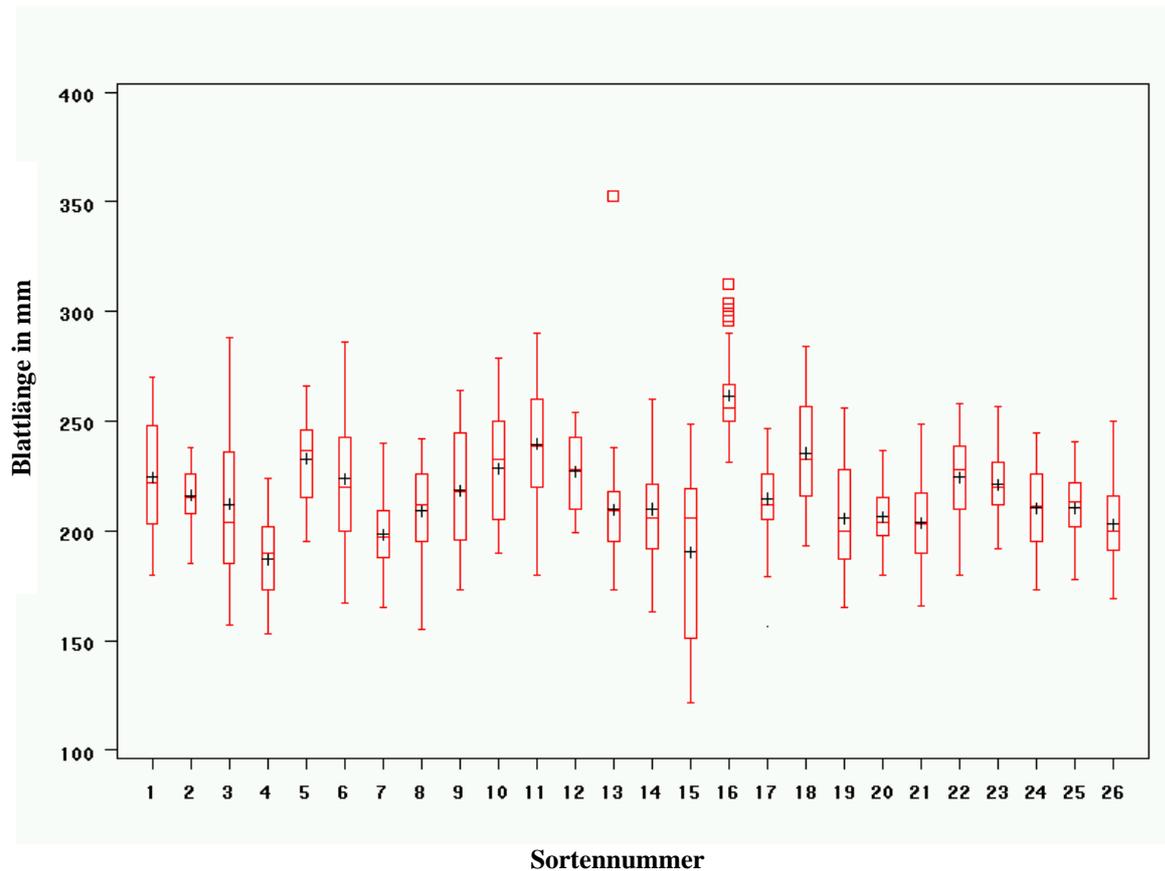


Abbildung 1. Box-Plot für die Blattlänge von 26 Sorten von Raps

2.2.6 Wenn abweichende Beobachtungen entdeckt werden, ist es wichtig zu versuchen herauszufinden, weshalb die Beobachtungen voneinander abweichen. In einzelnen Fällen ist es möglich, zum Feld zurückzukehren und zu überprüfen, ob die Parzelle durch externe Faktoren (z. B. Kaninchen) beschädigt wurde oder ein Meßfehler auftrat. Im letzteren Fall ist eine Korrektur möglich. In anderen Fällen kann es notwendig sein, die früheren Noten (oder andere Messungen aus derselben Pflanze/Parzelle) zu untersuchen, um den Grund für die abweichende Beobachtung herauszufinden. Im allgemeinen sollten die Erfassungen nur gestrichen werden, wenn gute Gründe hierfür vorliegen.

### **2.3 Annahmen für die statistische Analyse und Validierung dieser Annahmen**

Wenn die Daten statistisch analysiert werden sollen, dann müssen die Annahmen hinter der Theorie, auf die sich die statistischen Verfahren stützen – zumindest ungefähr – erfüllt werden. Dieser Abschnitt beschreibt die Annahmen hinter den üblichsten Verfahren der statistischen Analyse, die bei der DUS-Prüfung angewandt werden. Anschließend folgt ein Abschnitt über die Validierung der für die statistische Analyse erforderlichen Annahmen: Er beschreibt, wie sie bewertet werden können.

Die hier beschriebenen Verfahren für die Validierung der Annahmen hinter den statistischen Verfahren gelten für die Analysen von Einzelversuchen (randomisierte Blöcke). Die Grundsätze sind jedoch dieselben, wenn Daten aus mehreren Anbauversuchen über die Jahre analysiert werden. Anstelle der Parzellenmittelwerte werden die Analysen dann

aufgrund der Sortenmittelwerte pro Jahr durchgeführt, und die Blöcke entsprechen dann den Jahren.

### 2.3.1 Annahmen für die statistische Analyse, die mit der Varianzanalyse verbunden sind

#### *2.3.1.1 Einleitung*

2.3.1.1.1 Zunächst ist es wesentlich, daß der Anbauversuch angemessen angelegt wird und randomisiert ist. Die wichtigsten Annahmen der Varianzanalyseverfahren sind:

- unabhängige Erfassungen
- Varianzhomogenität
- Additivität der Block- und Sorteneffekte für eine randomisierte Blockanlage
- normalverteilte Erfassungen (Residuen)

2.3.1.1.2 Man könnte auch erwähnen, daß es bei den Daten keine Fehler geben sollte. Es ist jedoch nicht notwendig, dies als Annahme anzugeben, erstens, weil dies bereits im vorhergehenden Abschnitt über die Datenvalidierung erfaßt wird, und zweitens, weil die obigen Annahmen versagen würden, wenn es Fehler (oder zumindest große Fehler) gäbe, da die Erfassungen nicht normalverteilt sind und unterschiedliche Varianzen ausweisen (Nichthomogenität der Varianzen).

2.3.1.1.3 Die hier erwähnten Annahmen sind von höchster Bedeutung, wenn die auf der Methode der kleinsten Quadrate basierenden statistischen Verfahren für die Prüfung der Hypothesen angewandt werden. Wenn diese statistischen Verfahren nur für die Schätzung der Effekte (Mittelwerte) angewandt werden, sind die Annahmen weniger wichtig, und die Annahme der normalverteilten Erfassungen ist nicht notwendig.

#### *2.3.1.2 Unabhängige Erfassungen*

Dies ist eine äußerst wichtige Annahme. Sie bedeutet, daß keine Erfassungen auf anderen Erfassungen in derselben Analyse abhängen dürfen (die Abhängigkeit zwischen Erfassungen kann in das Modell eingebaut werden, wurde jedoch nicht in COYD und COYU oder in die übrigen in Dokument TGP/8 enthaltenen Verfahren eingebaut). Die Abhängigkeit kann beispielsweise durch Konkurrenz zwischen benachbarten Parzellen, fehlende Randomisierung oder ungeeignete Randomisierung verursacht werden. Weitere Einzelheiten über die Sicherstellung der Unabhängigkeit der Erfassungen sind in Teil I: Abschnitt 1.5.3.3.7 „Elemente des Anbauversuchs bei Einsatz der statistischen Analyse“ enthalten.

#### *2.3.1.3 Varianzhomogenität*

Die Varianzhomogenität bedeutet, daß die Varianz aller Erfassungen, abgesehen von der Zufallsvariation, identisch sein sollte. Typische Abweichungen von der Annahme der Varianzhomogenität fallen meistens in eine der folgenden zwei Gruppen:

- i) Die Varianz hängt vom Mittelwert ab. Je höher beispielsweise der Mittelwert ist, desto höher ist die Standardabweichung. In diesem Fall können die Daten häufig so transformiert werden, daß die Varianzen auf der transformierten Skala  $n$  ungefähr homogen sein können. Einige typische Transformationen von Merkmalen sind: die logarithmische Transformation (bei der die

Standardabweichung ungefähr proportional zum Mittelwert ist), die Quadratwurzeltransformation (bei der die Varianz ungefähr proportional zum Mittelwert ist, z. B. Zählungen) und die Winkeltransformation (bei der die Varianz an beiden Enden der Skala gering und zwischendurch höher ist, was typisch für Prozentsätze ist).

ii) Die Varianz hängt beispielsweise von der Sorte, dem Jahr oder dem Block ab. Wenn die Varianzen von solchen Variablen in einer Weise abhängen, die nicht mit dem Mittelwert verbunden ist, ist es nicht möglich, die Varianzhomogenität durch Transformation zu erreichen. In diesen Fällen könnte es notwendig sein, entweder höherentwickelte statistische Verfahren anzuwenden, die ungleiche Varianzen berücksichtigen können, oder die Gruppe von Erfassungen mit abweichenden Varianzen auszuschließen (wenn lediglich einige wenige Erfassungen abweichende Varianzen aufweisen). Zur Verdeutlichung der Bedeutung der Varianzheterogenität muß man sich einen Anbauversuch mit 10 Sorten vorstellen, bei dem die Sorten A, B, C, D, E, F, G und H je eine Varianz von 5 aufweisen, während die Sorten I und J je eine Varianz von 10 aufweisen. Die tatsächliche Wahrscheinlichkeit, Unterschiede zwischen diesen Sorten festzustellen, wenn sie faktisch denselben Mittelwert haben, wird in Tabelle 2 gezeigt. In Tabelle 2 beruhen die Sortenvergleiche auf der zusammengefaßten Varianz, wie dies bei der herkömmlichen ANOVA üblich ist. Wenn sie anhand des Signifikanzniveaus von 1 % verglichen werden, ist die Wahrscheinlichkeit, daß die beiden Sorten mit einer Varianz von 10 signifikant voneinander unterschiedlich werden, nahezu 5 Mal größer (4,6 %), als sie sein sollte. Andererseits nimmt die Wahrscheinlichkeit signifikanter Unterschiede zwischen zwei Sorten mit einer Varianz von 5 auf 0,5 % ab, wenn sie 1 % betragen sollte. Das bedeutet, daß es zu schwierig wird, Unterschiede zwischen zwei Sorten mit geringen Varianzen zu entdecken, und zu einfach, Unterschiede zwischen Sorten mit hohen Varianzen festzustellen.

Tabelle 2. Tatsächliche Wahrscheinlichkeit signifikanter Unterschiede zwischen zwei identischen Sorten, wenn die Varianzhomogenität angenommen, jedoch nicht erfüllt ist (die Sorten A bis H haben eine Varianz von 5, und die Sorten I und J haben eine Varianz von 10.)

Vergleiche, Sortennamen	Formaler Test des Signifikanzniveaus	
	1 %	5 %
A und B	0,5 %	3,2 %
A und I	2,1 %	8,0 %
I und J	4,6 %	12,9 %

### 2.3.1.4 Normalverteilte Erfassungen

Die Residuen sollten ungefähr normalverteilt sein. Das Residuum ist der Teil einer Erfassung, der nach der Erstellung eines Modells unerklärt bleibt. Es ist der Unterschied zwischen der Erfassung und der Prädiktion aus dem Modell. Die ideale Normalverteilung bedeutet, daß die Verteilung der Daten um den Mittelwert herum symmetrisch ist und die typische Glockenform aufweist (vergleiche Abbildung 2). Wenn die Residuen nicht ungefähr normalverteilt sind, kann das tatsächliche Signifikanzniveau vom Nominalniveau abweichen. Die Abweichung kann in beide Richtungen gehen, je nach der Art und Weise, wie die tatsächliche Verteilung der Residuen von der Normalverteilung abweicht. Die Abweichung von der Normalverteilung ist in der Regel jedoch nicht so bedeutend wie die Abweichungen von den vorherigen beiden Annahmen.

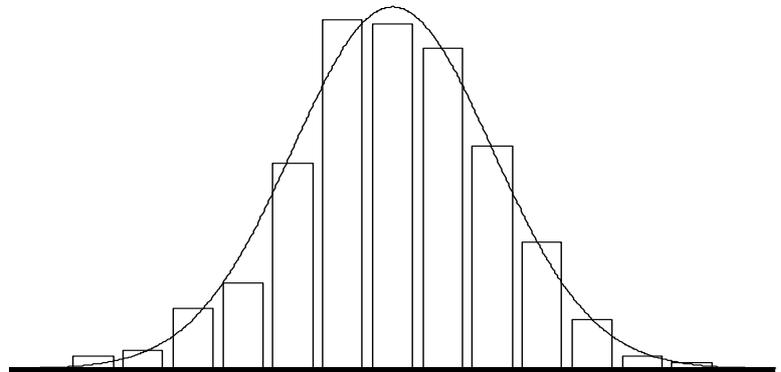


Abbildung 2. Histogramm für normalverteilte Daten; die ideale Normalverteilung ist als Kurve dargestellt

### 2.3.1.5 Additivität der Block- und Sorteneffekte

2.3.1.5.1 Die Effekte der Blöcke und Sorten werden als additiv angenommen, weil die Fehlergröße die Summe der Zufallsvariation und der Interaktion zwischen Block und Sorte ist. Das bedeutet, daß der Effekt einer gegebenen Sorte in allen Blöcken gleich ist. Dies wird in Tabelle 3 gezeigt, in der die Parzellenmittelwerte künstlicher Daten (die Blattlänge in mm) für zwei kleine Versuche mit drei Blöcken und vier Sorten angegeben werden. Im Versuch I sind die Effekte der Blöcke und Sorten additiv, weil die Unterschiede zwischen beliebigen zwei Sorten in allen Blöcken gleich sind; z. B. betragen die Unterschiede zwischen den Sorten A und B in allen drei Blöcken 4 mm. Im Versuch II sind die Effekte nicht additiv; z. B. betragen die Unterschiede zwischen den Sorten A und B in den drei Blöcken 2, 2 und 8 mm.

Tabelle 3. Künstliche Parzellenmittelwerte der Blattlänge in mm aus zwei Versuchen, die die additiven Block- und Sorteneffekte (links ) und die nichtadditiven Block- und Sorteneffekt (rechts) angeben.)

Versuch I				Versuch II			
Sorte	Block			Sorte	Block		
	1	2	3		1	2	3
A	240	242	239	A	240	242	239
B	244	246	243	B	242	244	247
C	245	247	244	C	246	244	243
D	241	243	240	D	241	242	241

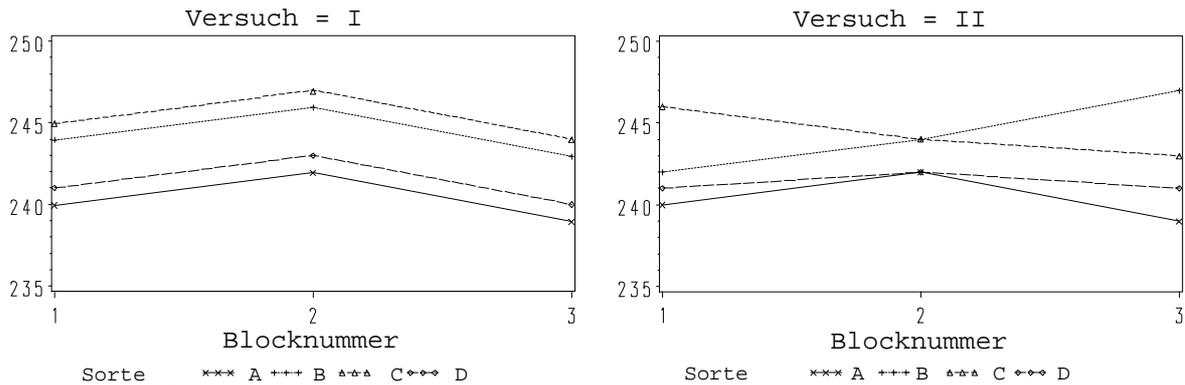


Abbildung 3. Künstliche Parzellenmittelwerte aus zwei Versuchen, die die additiven Block- und Sorteneffekte (links) und die nichtadditiven Block- und Sorteneffekt (rechts) anhand derselben Daten wie in Tabelle 3 angeben

2.3.1.5.2 In Abbildung 3 sind dieselben Daten graphisch dargestellt. Die graphische Darstellung der Mittelwerte gegenüber den Blocknummern und die Zusammenlegung der Erfassungen aus denselben Sorten durch gerade Linien erzeugen die Diagramme. Die graphische Darstellung der Mittelwerte gegenüber den Sortennamen und die Zusammenlegung der Erfassungen aus denselben Blöcken hätten ebenfalls verwendet werden können (und werden vielleicht bevorzugt, insbesondere wenn zahlreiche Sorten in derselben Abbildung gezeigt werden sollen). Die Annahme der Additivität ist erfüllt, wenn die Linien für die Sorten (abgesehen von der Zufallsvariation) parallel verlaufen. Da für jede Sorte in jedem Block lediglich ein einziger Datenwert vorliegt, ist es nicht möglich, die Interaktionseffekte und die Zufallsvariation voneinander zu trennen. Somit ist die Situation in der Praxis nicht so schön und klar wie hier, weil die Effekte durch die Zufallsvariation verschleiert werden können.

## 2.3.2 Validierung der Annahmen für die statistische Analyse

### 2.3.2.1 Einleitung

2.3.2.1.1 Der Hauptzweck der Validierung ist es zu überprüfen, ob die Annahmen, die den statistischen Analysen zugrundeliegen, erfüllt sind. Sie dient jedoch auch als Zweitkontrolle, ob die Daten fehlerfrei sind.

2.3.2.1.2 Es gibt verschiedene Verfahren, die bei der Validierung der Annahmen angewandt werden können. Einige von diesen sind:

- Durchsicht der Daten zur Überprüfung der Annahmen
- Erstellung graphischer Darstellungen oder Abbildungen zur Überprüfung der Annahmen
- Durchführung formaler statistischer Tests für die verschiedenen Arten von Annahmen. In der Literatur sind mehrere Verfahren für die Prüfung auf Ausreißer, Varianzhomogenität, Additivität und Normalverteilung zu finden. Diese Verfahren werden hier nicht erwähnt, teils weil viele von ihnen von Annahmen abhängen, die die Gültigkeit von COYD und COYU nicht nennenswert beeinträchtigen, und teils weil die Aussagekraft dieser Verfahren weitgehend von der Stichprobengröße abhängt (was bedeutet, daß ein schwerer Mangel an Annahmen bei kleinen Datensätzen unentdeckt bleiben kann, während kleine und unbedeutende Abweichungen in großen Datensätzen statistisch signifikant werden können)

### 2.3.2.2 *Datendurchsicht*

In der Praxis ist dieses Verfahren nur anwendbar, wenn einige wenige Erfassungen überprüft werden müssen. Für große Datensätze nimmt dieses Verfahren zuviel Zeit in Anspruch, ist beschwerlich und das Risiko, daß verdächtige Daten übersehen werden, steigt mit zunehmender Durchsicht der Daten an. Zudem ist es bei der Anwendung dieses Verfahrens sehr schwierig, die Verteilung der Daten und den Grad der Varianzhomogenität zu beurteilen.

### 2.3.2.3 *Erstellung von Abbildungen*

2.3.2.3.1 Es können verschieden Arten von Abbildungen erstellt werden, die für die verschiedenen zu validierenden Aspekte zweckdienlich sind. Viele davon bestehen aus der verschiedenartigen graphischen Darstellung der Residuen. (Die Residuen sind die Unterschiede zwischen den erfaßten Werten und den durch das statistische Modell prädiktierten Werte).

2.3.2.3.2 Die graphische Darstellung der Residuen gegenüber den prädiktierten Werten kann verwendet werden, um die Abhängigkeit der Varianz vom Mittelwert zu beurteilen. Wenn es keine Abhängigkeit gibt, sollten die Erfassungen ungefähr (ohne systematische Abweichung) in eine symmetrisch um null angeordnete horizontale Bande fallen (Abbildung 4). Wenn die Varianz mit dem Mittelwert zunimmt, fallen die Erfassungen ungefähr in einen Trichter mit dem schmalen Ende nach links gerichtet. Ausreißende Erfassungen, die Fehler sein können, werden in einer solchen Abbildung als Erfassungen gezeigt, die deutlich aus der durch die meisten anderen Erfassungen gebildeten horizontalen Bande ausgerissen sind. In dem Beispiel in Abbildung 4 scheinen keine Erfassungen Ausreißer zu sein (der Wert in der einen linken unteren Ecke, wo das Residuum etwa -40 mm beträgt, mag auf den ersten Blick so aussehen, doch haben mehrere Erfassungen positive Werte derselben numerischen Größe). Hier ist es wichtig anzumerken, daß ein Ausreißer nicht zwangsläufig ein Fehler ist und auch, daß ein Fehler nicht zwangsläufig als Ausreißer erscheint.

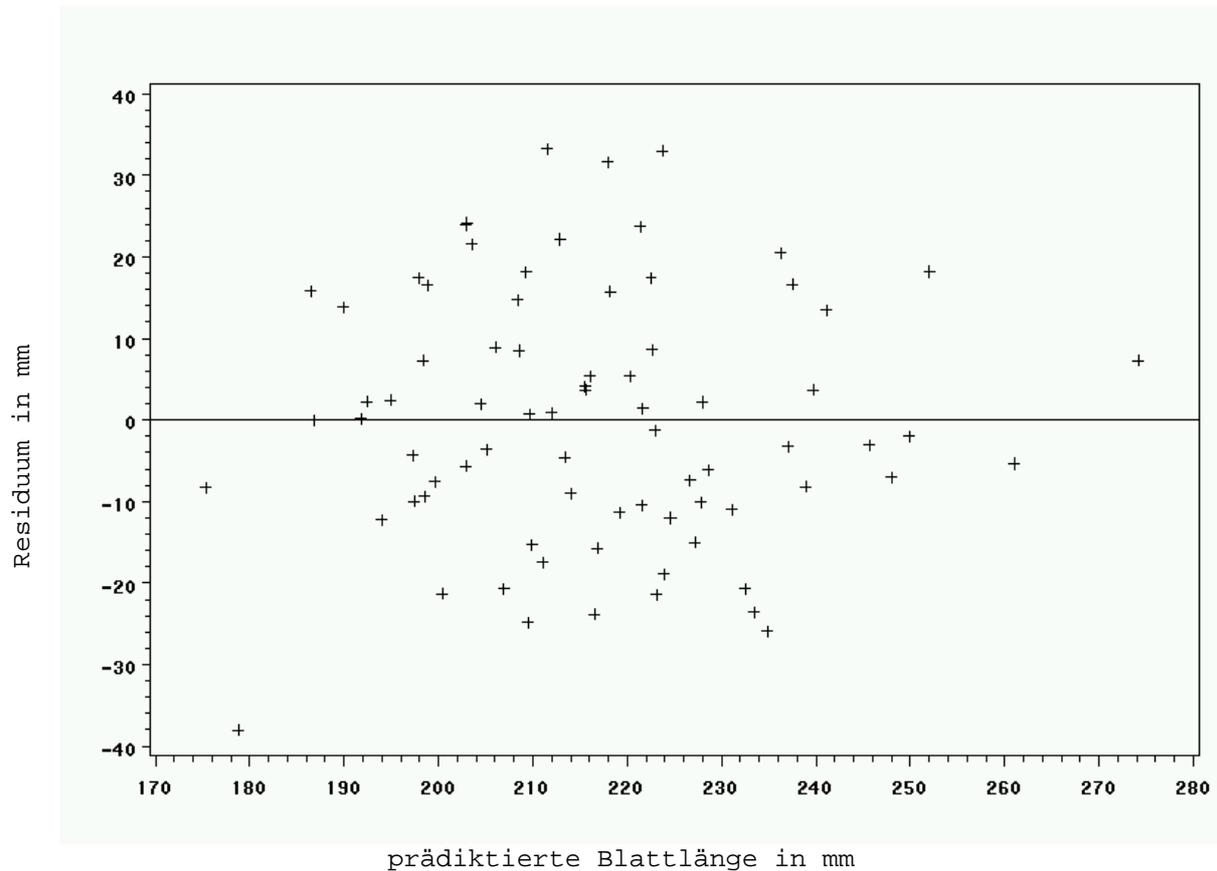


Abbildung 4. Graphische Darstellung der Residuen gegenüber den durch die graphische Darstellung prädiktierten Werten für die Blattlänge bei 26 Sorten von Raps in 3 Blöcken

2.3.2.3.3 Die Residuen können auch für ein Histogramm, wie Abbildung 2, verwendet werden, anhand dessen die Annahme bezüglich der Verteilung beurteilt werden kann.

2.3.2.3.4 Die Spannweite (Höchstwert minus Mindestwert) oder Standardabweichung für jede Parzelle kann gegenüber anderen Variablen wie den Parzellenmittelwerten, der Sortennummer oder der Parzellennummer graphisch dargestellt werden. Diese Abbildungen (Abbildung 5) können zweckdienlich sein, um Sorten mit extrem großer Variation zu ermitteln (alle Parzellen der Sorte mit hohem Wert) oder Parzellen, in denen die Variation extrem hoch ist (möglicherweise verursacht durch eine Einzelpflanze). Es ist deutlich zu ersehen, daß die Spannweite für eine der Parzellen der Sorten 13 weit höher als in den übrigen beiden Parzellen ist. Auch die Spannweite in einer der Parzellen der Sorte 3 scheint verhältnismäßig groß zu sein.

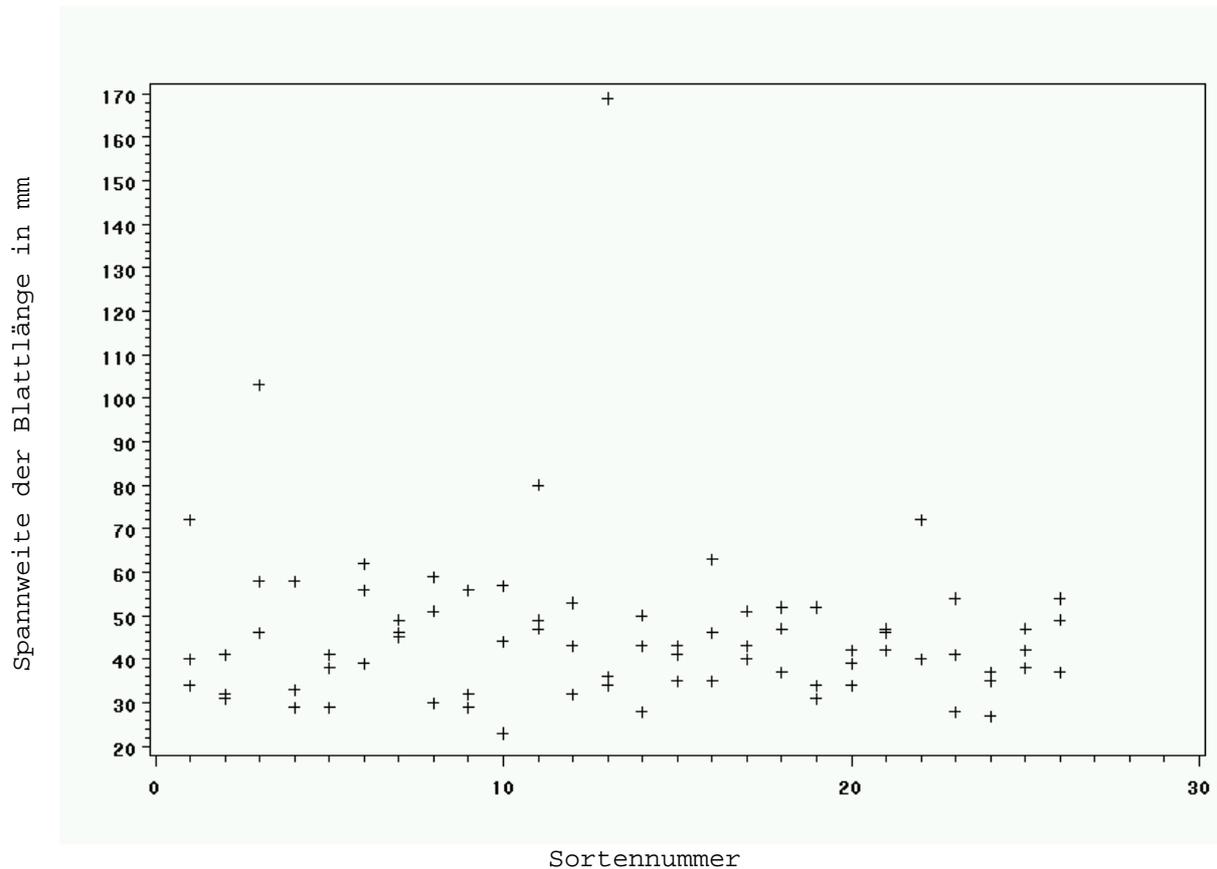
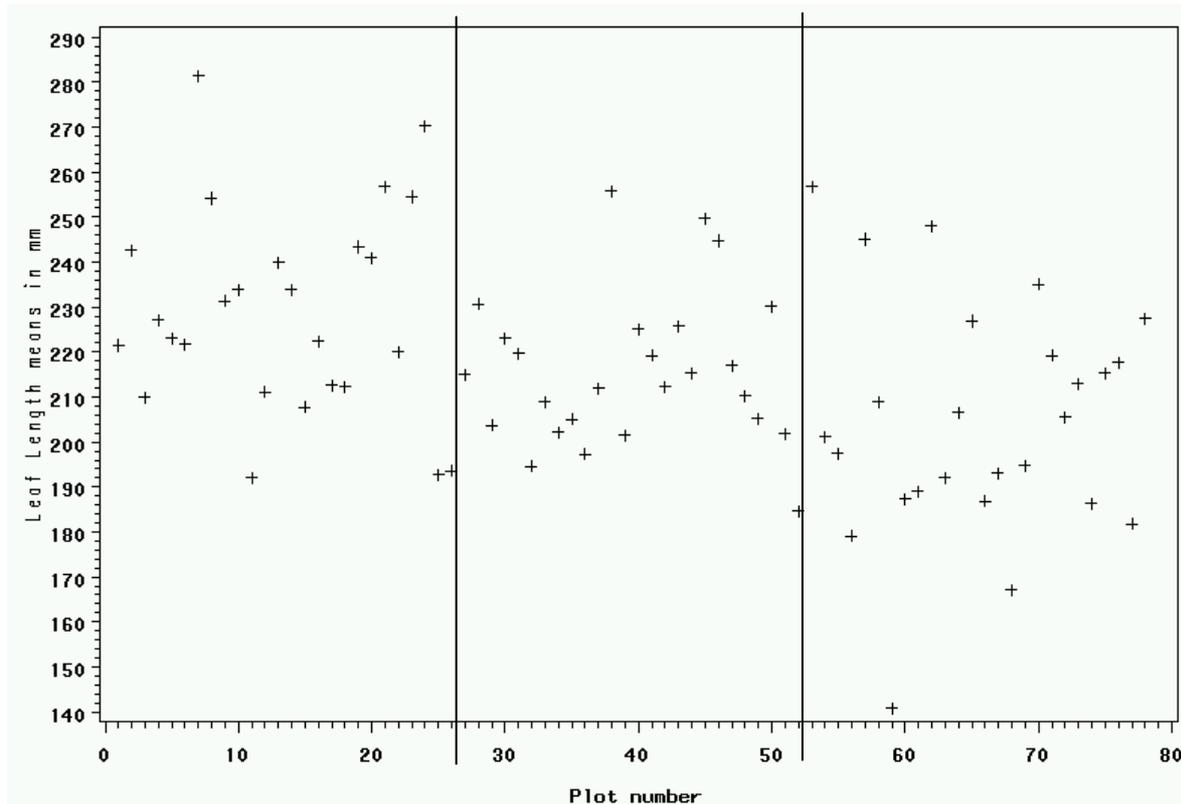


Abbildung 5. Unterschiede zwischen Minimum und Maximum der 20 Blattlängen für 3 Parzellen gegenüber der der Sortennummer von Raps

2.3.2.3.5 Eine Abbildung mit den Parzellenmittelwerten (oder bereinigten Sortenmittelwerten) gegenüber der Parzellennummer kann verwendet werden, um festzustellen, ob das Merkmal vom Prüfungsort im Feld abhängt (Abbildung 6). Dies setzt selbstverständlich voraus, daß die Parzellen so numeriert sind, daß die Nummern den relativen Prüfungsort angeben. Im Beispiel in Abbildung 6 ist ein deutlicher Trend vorhanden, der zeigt, daß die Blattlänge mit der Parzellennummer geringfügig abnimmt. Der Großteil des Trends über das für den Anbauversuch benutzte Gelände erklärt sich – in diesem Fall – jedoch durch die Unterschiede zwischen Blöcken (Parzelle 1-26 ist Block 1, Parzelle 27-52 ist Block 2 und Parzelle 53-78 ist Block 3).



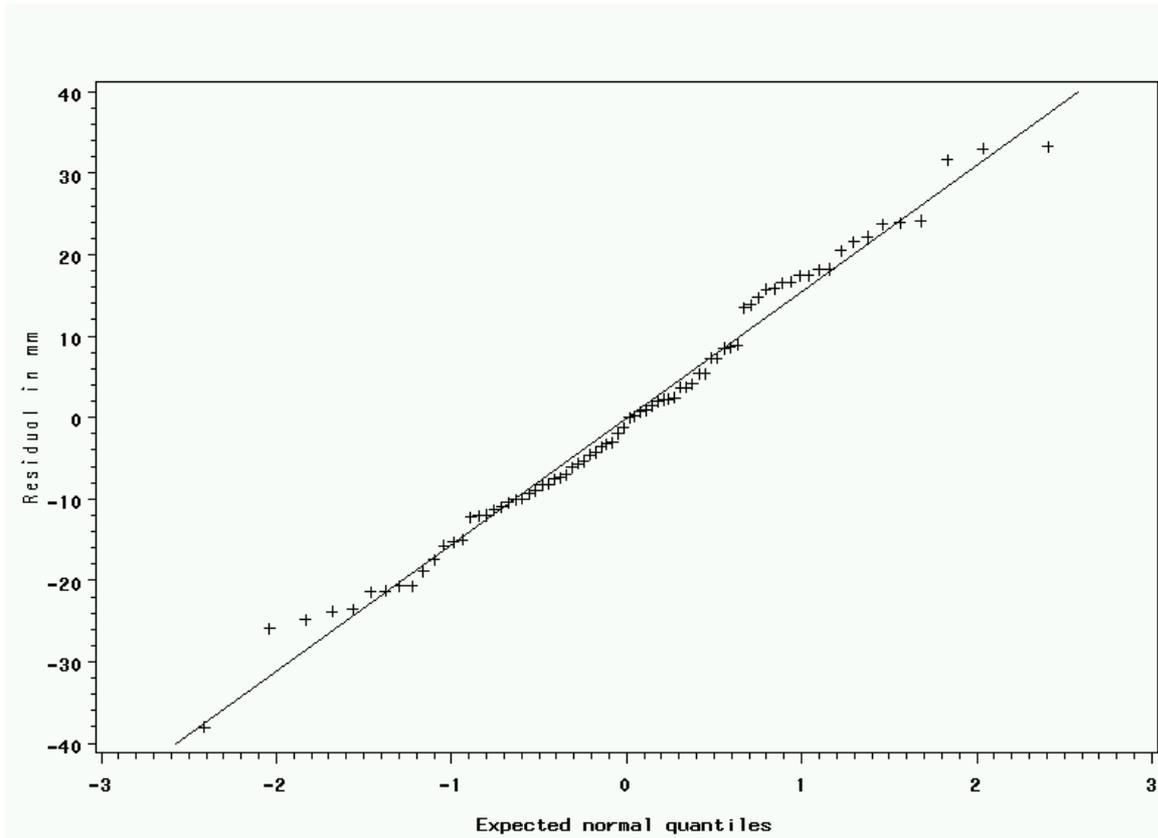
Mittelwerte der Blattlänge in mm

Parzellennummer

Abbildung 6. Parzellenmittelwerte von 20 Blattlängen gegenüber Parzellennummern

2.3.2.3.6 Die Parzellenmittelwerte können auch für die Erstellung einer Abbildung verwendet werden, in der die Additivität der Block- und Sorteneffekte visuell überprüft werden kann (vergleiche Abbildung 3).

2.3.2.3.7 Normalwahrscheinlichkeitsdiagramme (Abbildung 7). Diese Art Diagramm wird zur Beurteilung verwendet, inwieweit die Verteilung der Variable der Normalverteilung folgt. Die ausgewählte Variable wird in einem Streudiagramm gegenüber den Werten graphisch dargestellt, die „aus der Normalverteilung erwartet werden“. Das Standard-Normalwahrscheinlichkeitsdiagramm ist folgendermaßen aufgebaut. Zunächst werden die Residuen (Abweichungen von den Prädiktionen) nach Rangfolge geordnet. Von diesen Rängen berechnet das Programm die aus der Normalverteilung erwarteten Werte, nachstehend als z-Werte bezeichnet. Diese z-Werte sind auf der X-Achse des Diagramms eingezeichnet. Wenn die erfaßten Residuen (auf der Y-Achse eingezeichnet) normalverteilt sind, sollten alle Werte auf eine gerade Linie fallen. Wenn die Residuen nicht normalverteilt sind, weichen sie von der Linie ab. Ausreißer können auch in diesem Diagramm sichtbar werden. Wenn ein allgemeiner Mangel an Passung herrscht und die Daten ein klares Muster (z. B. eine S-Form) um die Linie herum zu bilden scheinen, muß die Variable möglicherweise in irgendeiner Weise transformiert werden.



Residuum in mm

Erwartete Normalquantile

Abbildung 7. Normalwahrscheinlichkeitsdiagramm für die Residuen der Blattlänge bei 26 Sorten von Raps in 3 Blöcken

### **3. WAHL DER STATISTISCHEN VERFAHREN FÜR DIE PRÜFUNG DER UNTERSCHIEDBARKEIT**

#### **3.1 Einleitung**

3.1.1 Dieser Abschnitt befaßt sich mit allgemeinen Überlegungen bei der Wahl geeigneter statistischer Verfahren für die Prüfung der Unterscheidbarkeit. Er enthält eine Erörterung der Faktoren, die die Wahl des Verfahrens beeinflussen, und da die von jedem Verfahren benutzte statistische Prüfung wesentlicher Teil dieses Verfahrens ist, enthält er eine kurze Erörterung der statistischen Tests, der Faktoren, die deren Wahl beeinflussen, und Kommentare zu deren Zweckdienlichkeit in bestimmten Situationen.

3.1.2 Statistische Verfahren werden am häufigsten für die Prüfung der Unterscheidbarkeit der gemessenen quantitativen Merkmale für fremdbefruchtende Sorten angewandt, wenn die Daten aus dem Anbauversuch für eine Sorte einer Variation unterworfen sind. Wegen dieser Variation sind auf statistischen Verfahren beruhende Unterscheidbarkeitskriterien erforderlich, um echte Sortenunterschiede von der Zufallsvariation zu unterscheiden und somit Entscheidungen darüber zu treffen, ob die Kandidatensorte mit einem gewissen Konfidenzniveau, daß die Entscheidung richtig ist, unterscheidbar ist.

3.1.3 Die Variation kann beispielsweise von Pflanze zu Pflanze, von Parzelle zu Parzelle und von Jahr zu Jahr auftreten. Ob eine einzige Wachstumsperiode oder mehr als eine einzige Wachstumsperiode erforderlich ist, um Gewißheit zu geben, daß die zwischen den Sorten erfaßten Unterschiede hinreichend stabil sind, hängt von dem an einer Art erfaßten Niveau oder vom Ausmaß der Variation aus diesen verschiedenen Quellen ab. Abschnitt 1.2 in Teil I dieses Dokuments erteilt Informationen über Wachstumsperioden.

#### **3.2 Statistische Verfahren bei zwei oder mehreren unabhängigen Wachstumsperioden**

##### 3.2.1 Einleitung

3.2.1.1 Verschiedene statistische Verfahren wurden entwickelt, um die Unterscheidbarkeit zu prüfen, wenn es mindestens zwei unabhängige Wachstumsperioden gibt. Die Wahl des Verfahrens hängt teils von der Pflanzenart und teils davon ab, ob die Anforderungen an den Anbauversuch und die Daten für die verschiedenen statistischen Verfahren erfüllt sind. Wenn diese Anforderungen nicht erfüllt sind, beispielsweise wenn nur eine oder sehr wenige bekannte Sorten für ein Taxon vorhanden sind und ein umfangreicher Anbauversuch daher nicht möglich ist, könnten andere geeignete Verfahren angewandt werden.

3.2.1.2 Die gemeinsamen Grundsätze der geeigneten statistischen Verfahren, die für die Prüfung der Unterscheidbarkeit angewandt werden, wenn es mindestens zwei unabhängige Wachstumsperioden gibt, sind:

- statistische Tests der Unterschiede zwischen Sortenmittelwerten werden durchgeführt, um zu bestimmen, ob die Unterschiede zwischen Sorten in bezug auf die Ausprägung ihrer Merkmale signifikant sind.

- die Voraussetzung, daß die Unterschiede in allen verschiedenen Wachstumsperioden stabil sein müssen. Diese Voraussetzung kann Teil des statistischen Tests wie bei der COYD-Methode, oder nicht Teil des statistischen Tests wie bei den 2x1 %- und Matchverfahren sein.

Um der Kürze willen wird im folgenden der Begriff ‚Jahr‘ benutzt, obwohl er für diese Zwecke mit dem Begriff ‚unabhängige Wachstumsperiode‘ austauschbar ist.

3.2.1.3 Beispiele für geeignete statistische Verfahren sind:

- a) Die Verfahren COYD und Langzeit-COYD für die Prüfung der Unterscheidbarkeit, die von der UPOV für die Analyse der Daten aus zwei oder mehreren Anbauversuchsjahren, wenn entweder mindestens eine gewisse Mindestzahl Sorten im Anbauversuch oder von Daten aus genügend Anbauversuchen in früheren Jahren vorliegen, entwickelt wurde. Ob die Unterschiede hinreichend stabil sind, wird anhand eines statistischen Tests aufgrund einer zweiseitigen LSD geprüft, um festzustellen, ob die Unterschiede bei den Sortenmittelwerten über die Jahre signifikant sind. Einzelheiten zu den Verfahren COYD und Langzeit-COYD und die Voraussetzungen für deren Anwendung sind in Dokument TGP/8 Teil II Abschnitt 3 enthalten.
- b) das 2x1 %-Verfahren zur Prüfung der Unterscheidbarkeit, das ebenfalls von der UPOV für die Analyse von Daten aus zwei oder mehreren Anbauversuchsjahren entwickelt wurde. Die Unterschiede werden in jedem Jahr anhand eines statistischen Tests aufgrund einer zweiseitigen LSD geprüft, um die Sortenmittelwerte innerhalb eines Jahres zu vergleichen. Ob die Unterschiede hinreichend stabil sind, wird anhand der Voraussetzung bestimmt, daß zwei Sorten in derselben Richtung auf dem Niveau 1 % in beiden Jahren oder, wenn die Anbauversuche in drei Jahren durchgeführt werden, in mindestens zwei der drei Jahre signifikant verschieden sind. Einzelheiten zum 2x1 %-Verfahren und ein Vergleich mit COYD sind in Dokument TGP/8 Teil II Abschnitt 4 enthalten.
- c) Der Match-Ansatz zur Prüfung der Unterscheidbarkeit wurde für die Analyse von Daten aus mehr als einer Wachstumsperiode entwickelt. Die Versuche werden in der ersten Wachstumsperiode vom Züchter und in der zweiten Wachstumsperiode von der Prüfungsbehörde durchgeführt (vergleiche Dokument TGP/6 „Organisation der DUS-Prüfung“, Abschnitt 2 „Beispiele für die Organisation der DUS-Prüfung“). Ob die Unterschiede hinreichend stabil sind, wird anhand eines statistischen Tests geprüft (z.B. LSD, Mehrfachreihentest (MRT), Chi-Quadrat oder exakter Fisher-Test) um zu beurteilen, ob die Unterschiede bei den Sortenmittelwerten im zweiten Jahr signifikant sind und mit der „Richtung der Unterschiede“ übereinstimmen, die von den Züchtern im ersten Jahr gemeldet wurden. Die Wahl des statistischen Tests hängt vom Ausprägungstyp des betreffenden Merkmals ab. Einzelheiten zum Match-Ansatz werden in Dokument TGP/8 Teil II, Abschnitt 7 dargelegt.

3.2.1.4 Im Zusammenhang mit Konsistenz und Harmonisierung sollte zur Kenntnis genommen werden, daß unterschiedliche statistische Verfahren zu unterschiedlichen Ergebnissen führen können.

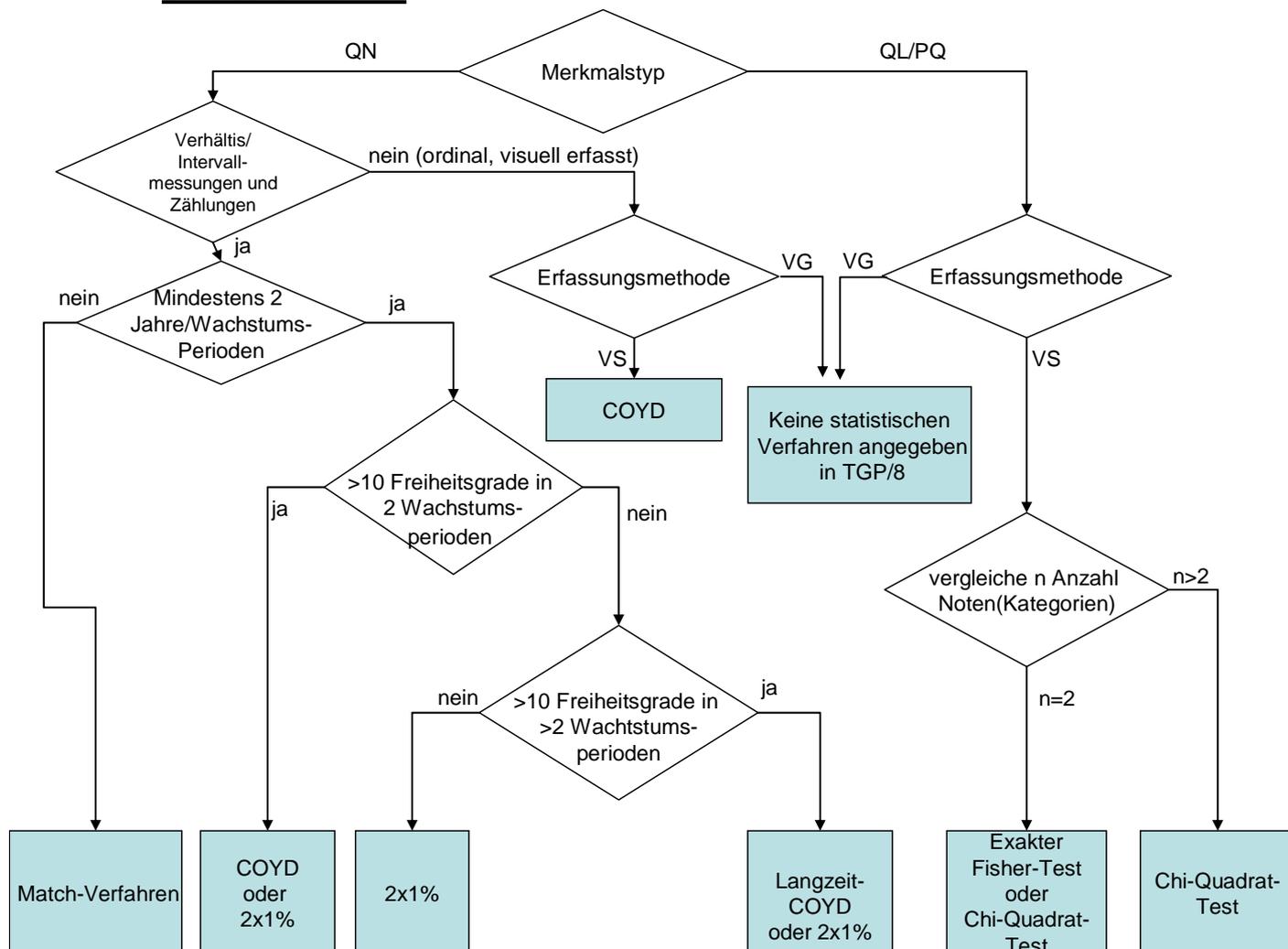
**3.3 Zusammenfassung der ausgewählten statistischen Verfahren für die Prüfung der Unterscheidbarkeit**

## 3.3.1 Ausgewählte Verfahren bei der DUS-Prüfung

	Erfassungsmethode	Mindestanzahl Jahre/Wachstumsperioden	Mindestanzahl Freiheitsgrade	zu prüfende Hypothese	Merkmaltyp	Sonstige
<b>COYD*</b>	MS/VS	2	20 in zwei Jahren/Wachstumsperioden	D/nicht D für Sortenmittelwerte	QN	-
<b>Langzeit-COYD*</b>	MS/VS	2	20 (Verwendung der Daten aus mehr als zwei Jahren/Wachstumsperioden)	D/nicht D für Sortenmittelwerte	QN	-
<b>2x1 %</b>	MS	2		D/nicht D für Sortenmittelwerte	QN	-
<b>Match-Verfahren</b>	VS	-	-	Hypothese für D aufgrund früher bekannter Fakten oder Grundsätze*	PQ/QN	2 oder mehrere Sorten durch ein Merkmal verglichen Ausprägungen zwei oder mehreren Kategorien zugeordnet Wert jeder Kategorie beträgt mehr als fünf
	VS	-	-	Hypothese für D aufgrund früher bekannter Fakten oder Grundsätze*	PQ/QN	2 Sorten durch ein Merkmal verglichen Ausprägungen zwei oder mehreren Kategorien zugeordnet Wert jeder Kategorie beträgt weniger als 10

\* Unter bestimmten Umständen könnten auch \*MG and VG anwendbar sein

**3.4 Voraussetzungen für die statistischen Verfahren für die Prüfung der Unterscheidbarkeit**



Andere statistische Verfahren, die nicht in obigem Diagramm enthalten sind, könnten für die Bestimmung der Unterscheidbarkeit geeignet sein.

[Teil II folgt]

## **TEIL II: AUSGEWÄHLTE VERFAHREN FÜR DIE DUS-PRÜFUNG**

### **1. DIE GAIA-METHODIK**

Die GAIA-Methodik wurde entwickelt, um die Anbauversuche zu optimieren, indem der Anbau von einigen Sorten aus der Sortensammlung vermieden wird. Der Grundsatz besteht darin, einen phänotypischen Abstand zwischen jedem Sortenpaar zu berechnen, wobei dieser Abstand die Summe der Abstände an jedem erfaßten Merkmal ist. Der Hintergrund dieses Verfahrens stützt sich auf die Möglichkeit, die dem Pflanzensachverständigen geboten wird, seine Erfahrung bezüglich der Zuverlässigkeit der erfaßten Unterschiede einzubringen, indem der Unterschied für jedes erfaßte Merkmal gewichtet wird.

Die GAIA-Methodik wird hauptsächlich nach einer ersten Wachstumsperiode angewandt, um diejenigen Sorten aus der Sortensammlung zu identifizieren, die von der(n) darauffolgenden Wachstumsperiode(n) ausgeschlossen werden können, weil sie von allen Kandidatensorten „unterscheidbar plus“ sind (vergleiche TGP/8/1 Teil II Abschnitt 1.3.2.1). GAIA kann auch die ähnlichen Sorten identifizieren, auf die der DUS-Prüfer seine besondere Aufmerksamkeit in der darauffolgenden Wachstumsperiode richten muß.

#### **1.1 Gründe für die Summierung und Gewichtung erfaßter Unterschiede**

1.1.1 Bei der Prüfung der Unterscheidbarkeit beobachtet ein DUS-Prüfer zunächst eine Sorte nach Merkmalen. Bei ähnlichen Sorten berücksichtigt der DUS-Prüfer auch die Gesamtheit aller erfaßten Unterschiede. Die GAIA-Software unterstützt den DUS-Prüfer bei der Prüfung der Unterschiede nach Merkmalen und für die Gesamtheit aller Merkmale.

1.1.2 Ein DUS-Prüfer stellt möglicherweise fest, daß zwei Sorten nach der ersten Wachstumsperiode so deutlich unterscheidbar sind, daß der Vergleich nicht wiederholt werden muß. Diese beiden Sorten, die „unterscheidbar plus“ (vergleiche TGP/8/1 Teil II Abschnitt 1.3.2.1) sind, sind deutlich unterscheidbar.

1.1.3 Ein DUS-Prüfer kann eine Situation haben, in der zwei Sorten verschiedene Noten erhalten, doch werden die beiden Sorten vom Prüfer als ähnlich angesehen. Der Unterschied könnte darauf zurückzuführen sein, daß die Sorten nicht nahe nebeneinander angebaut wurden (d. h. unterschiedliche Umweltbedingungen hatten), oder auf die Unbeständigkeit des Beobachters bei der Beurteilung der Noten usw.

1.1.4 Die Merkmale sind unterschiedlich anfällig für Umwelteinflüsse, und die Genauigkeit, mit der sie erfaßt werden (d. h. visuelle Erfassung/Messung), kann ebenfalls variieren. Für Merkmale, die umweltabhängig sind und nicht sehr genau erfaßt werden, benötigt der Prüfer einen großen Unterschied zwischen der Sorte A und der Sorte B, um sicherzugehen, daß der erfaßte Unterschied die Unterscheidbarkeit angibt.

1.1.5 Für Merkmale, die umweltunabhängig sind und die genau erfaßt werden, kann sich der Prüfer auf einen kleineren Unterschied zwischen der Sorte A und der Sorte B verlassen.

1.1.6 Bei der GAIA-Methode bestimmt der Prüfer für jedes erfaßte Merkmal die geeigneten Gewichtungen für die erfaßten Unterschiede. Die Software berechnet die Summe der Gewichtungen und gibt dem Pflanzenprüfer an, welche Sortenpaare „unterscheidbar plus“

sind und welche nicht. Der Prüfer kann sodann entscheiden, welche der Sorten aus der Sortensammlung von der (den) darauffolgenden Wachstumsperiode(n) ausgeschlossen werden können, weil sie bereits von allen Kandidatensorten deutlich unterscheidbar sind.

## **1.2 Berechnung des phänotypischen GAIA-Abstandes**

1.2.1 Der Grundsatz der GAIA-Methode besteht darin, einen phänotypischen Abstand zwischen zwei Sorten zu berechnen, der der Gesamtabstand zwischen einem Sortenpaar ist, der sich aus der Addition der Gewichtungen aller Merkmale ergibt. Der phänotypische GAIA-Abstand beträgt demnach:

$$dist(i, j) = \sum_{k=1, nchar} W_k(i, j)$$

wobei:

$dist(i, j)$  der berechnete Abstand zwischen Sorte  $i$  und Sorte  $j$  ist.

$k$  das  $k$ . Merkmal aus den  $nchar$  Merkmalen ist, die für die Berechnung ausgewählt werden.

$W_k(i, j)$  die Gewichtung des Merkmals  $k$  ist, die eine Funktion des zwischen der Sorte  $i$  und der Sorte  $j$  erfaßten Unterschieds für dieses Merkmal  $k$  ist.

$$W_k(i, j) = f(|OV_{ki} - OV_{kj}|)$$

wobei  $OV_{ki}$  der an Merkmal  $k$  erfaßte Wert für die Sorte  $i$  ist.

1.2.2 Detaillierte Informationen sind in Abschnitt 1.3 enthalten.

## **1.3 Detaillierte Informationen über die GAIA-Methodik**

### **1.3.1 Gewichtung der Merkmale**

1.3.1.1 Es ist wichtig, die Korrelation zwischen Merkmalen bei der Gewichtung zu berücksichtigen. Wenn zwei Merkmale miteinander in Verbindung stehen (z.B. Höhe der Pflanze mit Rispe, Höhe der Pflanze ohne Rispe), ist es ratsam, nur eines davon in GAIA zu verwenden, um eine Doppelgewichtung zu vermeiden.

1.3.1.2 Die Gewichtung wird bei einem gegebenen Merkmal als der Beitrag zum Gesamtabstand zwischen einem Sortenpaar definiert. Dieses System muß für jede Art kalibriert werden, um das Gewicht zu bestimmen, das jedem Unterschied gegeben werden kann, und um die Zuverlässigkeit jedes Merkmals in einer gegebenen Umgebung und für die betreffende genetische Variabilität zu beurteilen. Aus diesem Grund ist die Rolle des Pflanzensachverständigen von wesentlicher Bedeutung.

1.3.1.3 Die Gewichtung hängt von der Größe des Unterschieds und vom einzelnen Merkmal ab. Die Gewichtungen werden von den Pflanzensachverständigen aufgrund ihrer Fachkenntnisse über die Pflanzenart und eines „Versuch-und-Kontrolle“-Lernprozesses („try-and-check“) festgelegt (vergleiche Diagramm 3 am Schluß dieser Anlage). Der Sachverständige kann geringen Unterschieden eine Gewichtung von null zuordnen; somit

könnte der Gesamtabstand null betragen, selbst wenn zwei Sorten bei zahlreichen Merkmalen verschiedene erfaßte Werte aufweisen. Für einen gegebenen Unterschied wird jedem Sortenpaar für ein gegebenes Merkmal die gleiche Gewichtung gegeben.

1.3.1.4 Die Gewichtung sollte einfach und konsistent sein. Der Pflanzensachverständige kann die Gewichtungen für ein Merkmal beispielsweise nur mit ganzzahligen Werten angeben, d. h. 0, 1, 2, 3 (oder mehr).

Wenn dies der Fall ist,

- wird den erfaßten Unterschieden, die für dieses Merkmal vom Pflanzensachverständigen als möglicherweise durch Umwelteffekte oder mangelnde Genauigkeit bei der Messung verursacht angesehen werden, eine Gewichtung von 0 gegeben
- eine Gewichtung von 1 ist die Mindestgewichtung, die als Nicht-null-Abstand beitragen kann
- eine Gewichtung von 3 wird als etwa 3 Mal höher in bezug auf die Konfidenz oder den Abstand als eine Gewichtung von 1 angesehen.

1.3.1.5 Der Schwellenwert der „Unterscheidbarkeit plus“ wird als Wert definiert, für den die Summe der Unterschiede mit Nicht-null-Gewichtung groß genug ist, um eine zuverlässige, deutliche Unterscheidung sicherzustellen.

1.3.1.6 Das Diagramm 3 ist ein Datenflußdiagramm, das beschreibt, wie ein iterativer „Versuch-und-Kontrolle“-Prozeß genutzt werden kann, um schrittweise eine zufriedenstellende Serie von Gewichtungen für eine gegebene Pflanze zu erzielen.

1.3.1.7 Folgendes einfache Beispiel bei *Zea mays* zeigt die Berechnung des Abstandes zwischen zwei Sorten:

Beispiel: Bei dem Merkmal „Gewichtungsmatrix Form des Kolbens“, das auf einer Skala 1 bis 3 erfaßt wird, teilte der Pflanzensachverständige denjenigen Unterschieden eine Gewichtung zu, die er für signifikant hält:

Form des Kolbens:

- 1 = kegelförmig
- 2 = kegelförmig-zylindrisch
- 3 = zylindrisch

Vergleiche zwischen Unterschieden bei Noten und Gewichtung		
	unterschiedlich bei Noten	Gewichtung
kegelförmig (1) kontra kegelförmig (1)	0	0
kegelförmig (1) kontra kegelförmig-zylindrisch (2)	1	2
kegelförmig (1) kontra zylindrisch (3)	2	6
kegelförmig-zylindrisch (2) kontra kegelförmig-zylindrisch (2)	0	0
kegelförmig-zylindrisch (2) kontra zylindrisch (3)	1	2
zylindrisch (3) kontra zylindrisch (3)	0	0

Wenn der Pflanzensachverständige eine Sorte  $i$  mit kegelförmigem Kolben (Note 1) mit einer Sorte  $j$  mit zylindrischem Kolben (Note 3) vergleicht, teilt er eine Gewichtung von 6 zu usw. Die Gewichtungen werden in Form einer Gewichtungsmatrix zusammengefaßt:

Gewichtungsmatrix $i$				
		Sorte $i$		
		1	2	3
Sorte $j$	1	0	2	6
	2		0	2
	3			0

Wenn der Pflanzensachverständige eine Sorte  $i$  mit kegelförmigem Kolben (Note 1) mit einer Sorte  $j$  mit zylindrischem Kolben (Note 3) vergleicht, teilt er eine Gewichtung von 6 zu.

## 1.3.2 Beispiele für Anwendungen

### *1.3.2.1 Bestimmung der „Unterscheidbarkeit plus“*

1.3.2.1.1 Der Schwellenwert für den phänotypischen Abstand, der für die Eliminierung von Sorten aus dem Anbauversuch angewandt wird, wird als „Unterscheidbarkeit plus“ bezeichnet und vom Pflanzensachverständigen auf ein Niveau angesetzt, das höher als der für die Begründung der Unterscheidbarkeit benötigte Unterschied ist. Dies stellt sicher, daß alle Sortenpaare, die einen gleichen oder größeren Abstand als den Schwellenwert (Unterscheidbarkeit plus) haben, unterscheidbar wären, wenn sie in einem erneuten Anbauversuch angebaut würden.

1.3.2.1.2 Der Schwellenwert „Unterscheidbarkeit plus“ muß auf der mit den Sorten aus der Sortensammlung gewonnenen Erfahrung beruhen und das Risiko minimieren, daß in einem darauffolgenden Anbauversuch ein Sortenpaar ausgeschlossen wird, das im Feld weiter verglichen werden müßte.

### 1.3.2.2 *Weitere Beispiele für Anwendungen*

#### *Anwendung des phänotypischen Abstands in der ersten Wachstumsperiode*

1.3.2.2.1 Eine Pflanzenart mit einer großen Sortensammlung, die nur die Merkmale auf einer Skala 1 bis 9 benutzt; die GAIA-Methodik ermöglicht die Auswahl der Sorten, die in den Anbauversuch einbezogen werden sollen. Dies kann genutzt werden, um die Anbauversuche der ersten Wachstumsperiode sowie der darauffolgenden Wachstumsperioden zu planen.

1.3.2.2.2 Bei Pflanzenarten mit relativ wenigen Kandidatensorten und einer kleinen Sortensammlung, die es dem Pflanzensachverständigen ermöglicht, alle Kandidatensorten (z. B. eine landwirtschaftliche Pflanzenart) und die geeigneten Sorten aus der Sortensammlung in zwei oder drei aufeinanderfolgenden Wachstumsperioden anzubauen. Dieselben Sorten werden in den Wachstumsperioden 1, 2 und 3 in einer randomisierten Anlage angebaut. Die Software unterstützt die Identifikation der Paare mit geringem Abstand, um den Sachverständigen in die Lage zu versetzen, bei der Besichtigung des Feldes seine Aufmerksamkeit auf diese bestimmten Fälle zu richten.

#### *Anwendung des phänotypischen Abstands nach dem ersten Anbauversuch*

1.3.2.2.3 Nach einer Wachstumsperiode (z. B. bei der Prüfung einer Zierpflanzenart) sind die absoluten Daten und Abstandsberechnungen ein objektives Mittel, um die Entscheidung des Sachverständigen zu sichern, weil die Qualität der Erfassung und die Zuverlässigkeit der erfaßten Unterschiede im Gewichtungssystem berücksichtigt wurden. Wenn weitere Wachstumsperioden notwendig sind, bevor eine Entscheidung getroffen wird, unterstützt die Software bei der Ermittlung der Fälle, auf die sich der Sachverständige konzentrieren muß.

1.3.2.2.4 Wenn es viele Kandidatensorten und viele Sorten in der Sortensammlung gibt und große Variabilität in der Art vorhanden ist (z. B. einer Gemüseart wie *Capsicum*); einerseits gibt es bereits deutliche Unterschiede nach lediglich einer Periode, andererseits sind jedoch einige Sorten sehr ähnlich. Damit die Pflanzensachverständigen bei ihren Erfassungen effizienter sind, möchten sie „ähnliche“ Sorten nahe beieinander anbauen. Die Rohergebnisse und die Abstände helfen bei der Wahl der „ähnlichen“ Sorten und der Entscheidung über die Anlage des Anbauversuchs für die darauffolgende Wachstumsperiode.

1.3.2.2.5 Bei Pflanzenarten, in denen es viele ähnliche Sorten gibt, für die es übliche Praxis ist, Seite-an-Seite-Vergleiche vorzunehmen, kann GAIA angewandt werden, um die ähnlichen Sorten nach der ersten Wachstumsperiode zu identifizieren, insbesondere wenn die Anzahl Sorten in einem Anbauversuch zunimmt, was die Ermittlung aller Problemsituationen erschwert. Die Software kann dazu beitragen, die weniger offensichtlichen Fälle „nicht zu übersehen“.

1.3.2.2.6 Bei vegetativ vermehrten Zierpflanzenarten dauert die Prüfung eine oder zwei Wachstumsperioden: Nach der ersten Wachstumsperiode sind einige Sorten aus der Sortensammlung im Anbauversuch deutlich von allen Kandidatensorten verschieden und müssen nicht in die zweite Wachstumsperiode einbezogen werden. Wenn die Anzahl Sorten

groß ist, können die Rohdaten und der Abstand (die Abstände) den Sachverständigen bei der Ermittlung der Sorten aus der Sortensammlung unterstützen, für die die zweite Wachstumsperiode nicht notwendig ist.

### 1.3.3 Berechnung des phänotypischen GAIA-Abstandes

Der Grundsatz der GAIA-Methode besteht darin, einen phänotypischen Abstand zwischen zwei Sorten zu berechnen, der die Summe der Gewichtungen ist, die der Pflanzensachverständige den von ihm erfaßten Merkmalen zuweist.

Der phänotypische GAIA-Abstand ist:

$$dist(i, j) = \sum_{k=1, nchar} W_k(i, j)$$

wobei:

$dist(i, j)$  der berechnete Abstand zwischen Sorte  $i$  und Sorte  $j$  ist.

$k$  das  $k$ . Merkmal aus den  $nchar$  Merkmalen ist, die für die Berechnung ausgewählt werden.

$W_k(i, j)$  die Gewichtung des Merkmals  $k$  ist, die eine Funktion des zwischen der Sorte  $i$  und der Sorte  $j$  erfaßten Unterschieds für dieses Merkmal  $k$  ist.

$$W_k(i, j) = f(|OV_{ki} - OV_{kj}|)$$

wobei  $OV_{ki}$  der an Merkmal  $k$  erfaßte Wert für die Sorte  $i$  ist.

Diese Berechnung des phänotypischen Abstands ermöglicht es,

- zwei Sorten zu vergleichen,
- eine gegebene Sorte mit allen übrigen Sorten zu vergleichen,
- alle Kandidatensorten mit allen beobachteten [Kandidatensorten + in den Versuch aufgenommene Sorten aus der Sortensammlung zu vergleichen
- alle möglichen Paarkombinationen zu vergleichen.

### 1.3.4 GAIA-Software

1.3.4.1 Die GAIA-Software ermöglicht die Berechnung des phänotypischen Abstands anhand der UPOV-Merkmale der Prüfungsrichtlinie, die allein oder in Kombination verwendet werden können. Der Nutzer kann über die Art der Daten und die Art und Weise ihrer Verwendung entscheiden. Er kann alle verfügbaren Merkmale oder aber verschiedene Untermengen von Merkmalen wählen.

1.3.4.2 Die hauptsächliche Anwendung von GAIA betrifft die Festlegung eines Schwellenwertes „unterscheidbar plus“, der einer zuverlässigen und deutlichen Unterscheidung entspricht.

1.3.4.3 Es ist daran zu erinnern, daß nicht alle Unterschiede mit einer Null-Gewichtung zum Abstand beitragen. Zwei Sorten können bei einer Reihe erfaßter Merkmale verschiedene Noten haben und schließlich einen Nullabstand aufweisen.

1.3.4.4 Nicht-null-Gewichtungen werden im Abstand summiert. Wenn der Abstand kleiner als der Schwellenwert „unterscheidbar plus“ ist, selbst wenn einige deutliche Unterschiede bei Noten oder Messungen vorhanden sind, werden die Sorten nicht als zuverlässig und deutlich unterscheidbar angegeben. Ist der Abstand größer als der vom Pflanzensachverständigen festgelegte Schwellenwert „unterscheidbar plus“, entspricht dies einem Fall, in dem einem Paarvergleich in einem weiteren Anbauversuch unnötig ist.

1.3.4.5 GAIA ermöglicht es dem Pflanzensachverständigen, die Schwellenwertparameter auf zwei weitere Arten für andere brauchbare Mittel als die Unterscheidbarkeit plus zu verwenden:

- Ein geringer Schwellenwert trägt dazu bei, die schwierigeren Fälle zu ermitteln (ähnliche oder verwandte Sorten zu identifizieren), auf die sich der Sachverständige im darauffolgenden Periode konzentrieren muß
- ein sehr hoher Schwellenwert ermöglicht die Visualisierung aller verfügbaren Rohdaten und der Gewichtungen für jedes Merkmal auf Bildschirmen und Ausdrucken

1.3.4.6 In der Praxis können verschiedene Schwellenwerte je nach verschiedenen Erfordernissen verwendet werden. Sie können vor einem Vergleich problemlos ausgewählt werden. Verschiedene Vergleiche können berechnet, gespeichert und mit ihren entsprechenden Schwellenwerten, der Merkmalsserie, der Sortenserie usw. aus der Datenbank abgerufen werden.

1.3.4.7 Die Software erstellt einen umfassenden Bericht für jeden Paarvergleich sowie eine Klassifikation aller Paarvergleiche von den unterschiedlicheren bis zu den ähnlicheren Paaren. Die Software berechnet einen Gesamtabstand, gibt jedoch auch alle individuellen absoluten Werte und den Beitrag jedes Merkmals zum Abstand an.

1.3.4.8 Zur Verkürzung der Berechnungszeit geht die Software, sobald der Schwellenwert für einen Vergleich zwischen zwei gegebenen Sorten erreicht ist, zum nächsten Sortenpaar weiter. Die restlichen Merkmale und ihre Rohwerte werden in der zusammenfassenden Datenausgabe nicht angezeigt und tragen nicht zum Abstand bei.

1.3.4.9 Abschnitt 1.3.5 zeigt eine Bildschirmskopie eines Anzeigebaums, aus der hervorgeht, wie der Sachverständige navigieren und die Ergebnisse der Berechnungen optisch anzeigen kann.

1.3.4.10 Die GAIA-Software wurde mit WINDEV entwickelt. Die allgemeinen Informationen (Art, Merkmale, Gewichtung usw.), die über die Sorten beschafften Daten und die Ergebnisse der Berechnungen werden in einer integrierten Datenbank gespeichert. Import- und Exportfunktionen ermöglichen die Nutzung anderer Informationssysteme in Verbindung mit der GAIA-Software. ODBC ermöglicht den gleichzeitigen Zugang zur GAIA-Datenbank und anderen Datenbanken.

1.3.4.11 Je Sorte können eine oder zwei Noten verwendet werden. Wenn eine Periode verfügbar ist, tritt eine Note auf. Zwei Noten sind beispielsweise vorhanden, wenn in einem gegebenen Jahr zwei Anbauversuche an verschiedenen Prüfungsorten durchgeführt werden oder wenn zwei Perioden am selben Prüfungsort erlangt werden. Für Elektrophoresedaten kann nur eine Beschreibung je Sorte eingegeben werden. Für Messungen sind mindestens zwei Werte (verschiedene Anbauversuche, Wiederholungen usw.) erforderlich, und der Nutzer kann wählen, welchen er bei der Berechnung benutzen will.

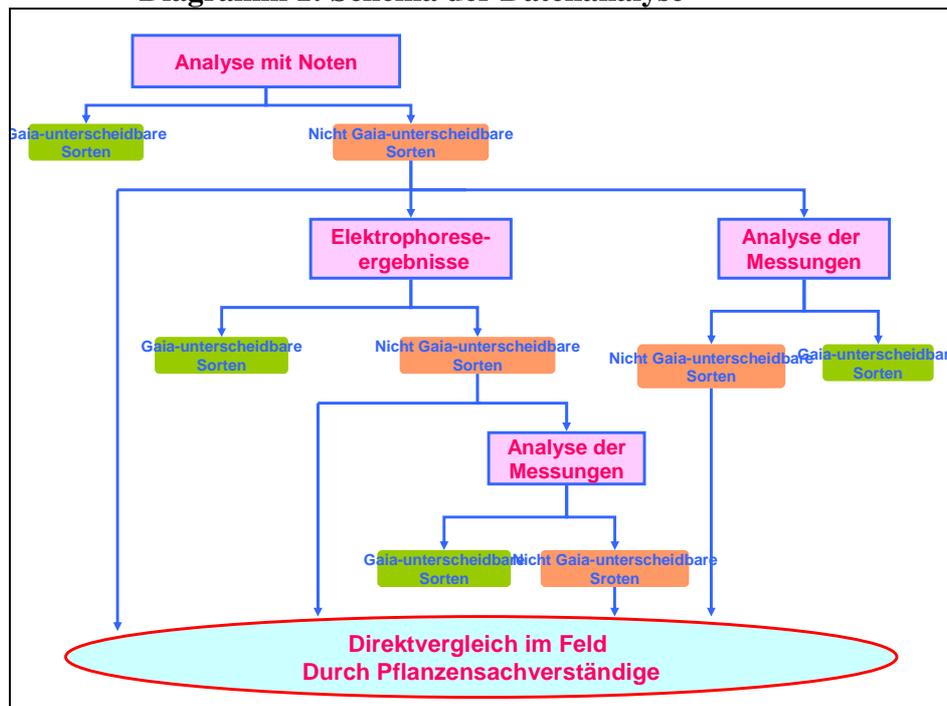
1.3.4.12 GAIA ist am geeignetsten für selbstbefruchtende und vegetativ vermehrte Sorten, kann jedoch auch für andere Sortentypen angewandt werden.

### 1.3.5 Beispiel mit Daten von *Zea mays*

#### 1.3.5.1 Einleitung

Die Software kann Noten, Messungen und/oder Elektrophoreseergebnisse verwenden. Diese Datentypen können allein oder in Kombination verwendet werden, wie in Diagramm 1 gezeigt.

**Diagramm 1: Schema der Datenanalyse**



In diesem Beispiel wird angenommen, daß der Pflanzensachverständige entschieden hat, einen Schwellenwert „Unterscheidbarkeit plus“  $S_{dist}$  von 10 zu benutzen.

#### 1.3.5.2 Analyse der Noten

1.3.5.2.1 Bei der qualitativen Analyse werden Noten (1 bis 9) verwendet. Die Noten können aus qualitativen, quantitativen und pseudoquantitativen Merkmalen stammen.

1.3.5.2.2 Für jedes Merkmal werden die Gewichtungen je nach Unterschieden zwischen Ausprägungsniveaus in einer Abstandsmatrix prädefiniert.

1.3.5.2.3 „Form des Kolbens“: auf einer Skala 1 bis 3 erfaßt; der Pflanzensachverständige wies den Unterschieden, die für signifikant angesehen werden, Gewichtungen zu, die größer als null sind:

1 = kegelförmig  
2 = kegelförmig-zylindrisch  
3 = zylindrisch

		Sorte ‚i‘		
		1	2	3
Sorte ‚j‘	1	0	2	6
	2		0	2
	3			0

1.3.5.2.4 Wenn der Pflanzensachverständige eine Sorte ‚i‘ mit kegelförmigem Kolben (Note 1) mit einer Sorte ‚j‘ mit zylindrischem Kolben (Note 3) vergleicht, wird eine Gewichtung von 6 zugeordnet.

1.3.5.2.5 „Länge der Lieschen“, auf einer Skala von 1 bis 9 erfaßt; der Pflanzensachverständige legte folgende Gewichtungsmatrix fest:

1 = sehr kurz  
2 = sehr kurz bis kurz  
3 = kurz  
4 = kurz bis mittel  
5 = mittel  
6 = mittel bis lang  
7 = lang  
8 = lang bis sehr lang  
9 = sehr lang

		Sorte ‚i‘								
		1	2	3	4	5	6	7	8	9
Sorte ‚j‘	1	0	0	0	2	2	2	2	2	2
	2		0	0	0	2	2	2	2	2
	3			0	0	0	2	2	2	2
	4				0	0	0	2	2	2
	5					0	0	0	2	2
	6						0	0	0	2
	7							0	0	0
	8								0	0
	9									0

1.3.5.2.6 Die Gewichtung zwischen einer Sorte ‚i‘ mit sehr kurzen Lieschen (Note 1) und einer Sorte ‚j‘ mit kurzen Lieschen (Note 3) ist 0. Der Sachverständige betrachtet einen Unterschied von drei Noten als Mindestunterschied für die Erkennung eines Nicht-null-Abstands zwischen zwei Sorten. Selbst wenn der Unterschied bei den Noten größer als 3 ist, behält der Sachverständige die Abstandsgewichtung bei 2, während bei sehr zuverlässigen Merkmalen einem Unterschied von 1 eine Gewichtung von 6 zugeordnet wird.

1.3.5.2.7 Der Grund für die Zuweisung einer geringeren Gewichtung für einige Merkmale im Vergleich zu anderen kann sein, daß sie weniger „zuverlässig“ oder „stabil“ (z. B. dem Umwelteffekt stärker unterworfen) sind und/oder so angesehen werden, daß sie einen geringeren Abstand zwischen den Sorten angeben.

1.3.5.2.8 Die Matrix für eine qualitative Analyse für fünf Merkmale für die Sorten A und B:

	Form des Kolbens	Länge der Lieschen	Typ des Korns	Anzahl Kornreihen	Durchmesser des Kolbens	
Noten für Sorte A (Skala 1 bis 9)	1	1	4	6	5	
Noten für Sorte B (Skala 1 bis 9)	3	3	4	4	6	
erfaßter Unterschied	2	2	0	2	1	
<i>Gewichtung gemäß dem Pflanzensachverständigen</i>	6	0	0	2	0	$D_{qual} = 8$

In diesem Beispiel  $D_{qual} = 8 < 10$  ( $S_{dist} = 10$  in diesen Beispiel) werden die Sorten A und B aufgrund dieser fünf Merkmale für „NICHT GAIA-unterscheidbar“ erklärt.

1.3.5.3 Elektrophoreseanalyse

1.3.5.3.1 In einigen UPOV-Prüfungsrichtlinien können Elektrophoreseergebnisse verwendet werden, wie bei *Zea mays*. Die Software läßt die Verwendung heterozygoter Loci nicht zu, sondern, im Einklang mit den Prüfungsrichtlinien, lediglich von homozygoten Loci. Die verwendeten Ergebnisse sind 0 (fehlend) und 1 (vorhanden) sowie die Kenntnis der Chromosomenzahl.

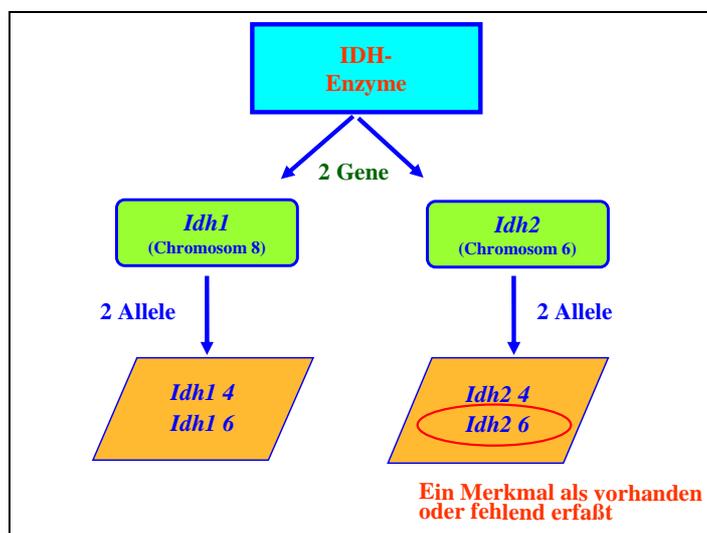


Diagramm 2: Das Enzym Isocitrat-Dehydrogenase (IDH) hat zwei Gene (*Idh1* und *Idh2*), die sich auf zwei verschiedenen Chromosomen befinden. Jedes davon hat zwei Allele, die als 1 (vorhanden) oder 0 (fehlend) erfaßt werden.

1.3.5.3.2 Elektrophoreseergebnisse werden als 0 oder 1 (fehlend oder vorhanden) benotet. Die Entscheidungsregel für die Zuweisung einer Gewichtung an zwei Sorten ist die Addition der Gewichtungszahl der erfaßten Unterschiede und der Gewichtungszahl der Chromosomen, die mit diesen Unterschieden verbunden sind (vergleiche nachstehendes Beispiel):

	Chromosom 8		Chromosom 6	
	<i>Idh1 4</i>	<i>Idh1 6</i>	<i>Idh2 4</i>	<i>Idh2 6</i>
Sorte A	0	1	1	0
Sorte B	0	1	0	1
Unterschied	0	0	1	1

1.3.5.3.3 In diesem Beispiel sind die Sorten A und B für vier Elektrophoreseergebnisse beschrieben:

*Idh1 4*, *Idh1 6*, *Idh2 4* und *Idh2 6*. Die Software prüft die Unterschiede und gibt den phänotypischen Abstand anhand der nachstehenden Berechnung an:

$$D_{\text{elec}} = 2 \times 0,25 + 1 \times 1 = 1,5$$

The diagram illustrates the calculation of  $D_{\text{elec}}$  with four text boxes explaining the components:

- 2** is the number of detected differences.
- 0,25** is the weight assigned by experts to the number of differences.
- 1** is the number of chromosomes in which differences were detected.
- 1** is the weight assigned by experts to the chromosome.

1.3.5.3.4 Diese Formel, die möglicherweise schwer verständlich ist, wurde von den Pflanzensachverständigen in Zusammenarbeit mit biochemischen Sachverständigen festgelegt. Sowohl die *Anzahl Unterschiede* als auch die *Anzahl Chromosomen, an denen Unterschiede erfaßt werden*, werden benutzt. Somit wird den Unterschieden weniger Bedeutung beigemessen, wenn sie auf demselben Chromosom vorkommen, als wenn sie auf verschiedenen Chromosomen vorkommen.

1.3.5.3.5 Nach der qualitativen and elektrophoretischen Analyse ist der phänotypische Abstand zwischen den Sorten A und B gleich:

$$D = D_{\text{qual}} + D_{\text{elec}} = 8 + 1,5 = 9,5$$

1.3.5.3.6 Der phänotypische Abstand ist *geringer als*  $S_{\text{dist}}$  ( $S_{\text{dist}}=10$  in diesem Beispiel); *deshalb werden die Sorten A und B als „NICHT GAIA-unterscheidbar“* angesehen.

1.3.5.3.7 Der Pflanzensachverständige kann entscheiden, ob er die Unterscheidbarkeit nicht allein aufgrund der Elektrophoreseanalyse begründen will. Bei der qualitativen Analyse ist ein minimaler phänotypischer Abstand erforderlich, um die Elektrophoreseergebnisse zu berücksichtigen. Dieser minimale phänotypische Abstand muß ebenfalls vom Pflanzensachverständigen festgelegt werden.

### 1.3.5.4 Analyse der Messungen

1.3.5.4.1 Die Analyse der Messungen berechnet die Unterschiede an erfaßten oder berechneten Messungen; Zählungen werden als Messungen behandelt.

1.3.5.4.2 Für jedes gemessene Merkmal wird der Vergleich von zwei Sorten durchgeführt, indem bei mindestens zwei verschiedenen Versuchseinheiten nach stabilen Unterschieden gesucht wird. Die Versuchseinheiten werden vom Nutzer je nach den in der Datenbank vorhandenen Daten festgelegt. Dies können beispielsweise Daten aus zwei geographischen Prüfungsorten der ersten Wachstumsperiode oder, im Fall eines einzigen geographischen Prüfungsorts, zwei oder drei Wiederholungen aus demselben Anbauversuch oder Daten aus zwei Perioden am selben Prüfungsort sein.

1.3.5.4.3 Damit ein Vergleich durchgeführt werden kann, müssen in denselben Versuchseinheiten zwei Sorten vorhanden sein. Die erfaßten Unterschiede müssen größer als einer der vom Pflanzensachverständigen festgelegten zwei Schwellenwerte (oder Mindestabstände) sein.

- $D_{\min\text{-inf}}$  ist der geringere Wert, aus dem eine Gewichtung zugewiesen wird,
- $D_{\min\text{-sup}}$  ist der höhere Mindestabstand. Diese Werte könnten willkürlich gewählt oder berechnet werden (15 % und 20 % des Mittelwerts für den Anbauversuch, oder LSD bei 1 % und 5 % usw.)

Für jeden Mindestabstand wird eine Gewichtung zugewiesen:

- $D_{\min\text{-inf}}$ : eine Gewichtung  $P_{\min}$  wird zugewiesen;
- $D_{\min\text{-sup}}$ : eine Gewichtung  $P_{\max}$  wird zugewiesen;
- der erfaßte Unterschied ist geringer als  $D_{\min\text{-inf}}$ : eine Gewichtung von null wird assoziiert.

1.3.5.4.4 Die Sorten A und B wurden für die Merkmale „Breite der Spreite“ und „Länge der Pflanze“ in zwei Anbauversuchen gemessen.

Für jeden Anbauversuch und jedes Merkmal entschied der Pflanzensachverständige,  $D_{\min\text{-inf}}$  und  $D_{\min\text{-sup}}$  festzulegen, indem die 15 % bzw. die 20 % des Mittelwerts für den Anbauversuch berechnet werden:

	Breite der Spreite		Länge der Pflanze	
	Anbauversuch 1	Anbauversuch 2	Anbauversuch 1	Anbauversuch 2
$D_{\min\text{-inf}} = 15\%$ des Mittelwerts des Anbauversuchs	1,2 cm	1,4 cm	28 cm	24 cm
$D_{\min\text{-sup}} = 20\%$ des Mittelwerts des Anbauversuchs	1,6 cm	1,9 cm	37 cm	32 cm

Für jedes Merkmal wies der Pflanzensachverständige folgende Gewichtung zu:

Eine Gewichtung  $P_{\min} = 3$  wird zugewiesen, wenn der Unterschied größer als  $D_{\min-\text{inf}}$  ist

Eine Gewichtung  $P_{\max} = 6$  wird zugewiesen, wenn der Unterschied größer als  $D_{\min-\text{sup}}$  ist

	Breite der Spreite		Länge der Pflanze		
	Anbau- versuch 1	Anbau- versuch 2	Anbau- versuch 1	Anbau- versuch 2	
Sorte A	9,9 cm	9,8 cm	176 cm	190 cm	
Sorte B	9,6 cm	8,7 cm	140 cm	152 cm	
Unterschied	0,3 cm	1,1 cm	36 cm	38 cm	
Gewichtung gemäß dem Pflanzen- sachverständigen	0	0	3	6	$D_{\text{quan}} = ?$

1.3.5.4.5 In diesem Beispiel sind die für das Merkmal „Breite der Spreite“ erfaßten Unterschiede geringer als  $D_{\min-\text{inf}}$ , so daß keine Gewichtung assoziiert wird. Andererseits ist ein Unterschied für das Merkmal „Länge der Pflanze“ größer als der Wert  $D_{\min-\text{inf}}$  und der andere größer als der Wert  $D_{\min-\text{sup}}$ . Diesen beiden Unterschieden werden verschiedene Gewichtungen zugewiesen.

1.3.5.4.6 Der Nutzer muß entscheiden, welche Gewichtung er für die Analyse benutzen will:

- die gewählte Gewichtung ist die dem geringsten Unterschied zugewiesene (minimalistische Option);
- die gewählte Gewichtung ist die dem größten Unterschied zugewiesene (maximalistische Option);
- Mittelwertoption: die gewählte Gewichtung ist die dem Mittelwert der übrigen zugewiesene (Mittelwertoption).

1.3.5.4.7 In diesem Beispiel entschied der Pflanzensachverständige, die geringste der beiden Gewichtungen zu wählen, so daß der auf Messungen beruhende phänotypische Abstand  $D_{\text{quan}} = 3$  ist.

1.3.5.4.8 Zusammenfassend ausgedrückt, ist der phänotypische Abstand zwischen den Sorten A und B nach allen Analysen :

$$D = D_{\text{qual}} + D_{\text{elec}} + D_{\text{quan}} = 8 + 1,5 + 3 = 12,5 > S_{\text{dist}}$$

1.3.5.4.9 Der phänotypische Abstand ist größer als der Unterscheidungsschwellenwert  $S_{\text{dist}}$ , der vom Pflanzensachverständigen auf 10 angesetzt wurde, so daß die Sorten A und B für „GAIA-unterscheidbar“ erklärt werden.

1.3.5.4.10 In diesem Beispiel „bestätigt“ die Anwendung der Elektrophoresedaten einen Abstand zwischen den beiden Sorten; aufgrund der qualitativen und quantitativen Daten allein wird der Schwellenwert jedoch überstiegen ( $8 + 3 = 11$  ist größer als 10).

1.3.5.4.11 Wenn der Schwellenwert auf 6 angesetzt worden wäre, wäre der Unterschied beim Merkmal „Form des Kolbens“ ausreichend gewesen, da die Sorte A kegelförmig und die Sorte B zylindrisch ist, was bereits ein deutlicher Unterschied ist.

1 = kegelförmig  
2 = kegelförmig-zylindrisch  
3 = zylindrisch

Sorte i			
	1	2	3
1	0	2	6
2		0	2
3			0

### 1.3.5.5 Messungen und Skala 1 bis 9 beim selben Merkmal

1.3.5.5.1 Für einige Pflanzen ist es die übliche Praxis, Werte aus Messungen auf einer Skala von 1 bis 9 zu erzeugen. Der Transformationsprozeß ist sehr einfach, bisweilen auch komplex.

1.3.5.5.2 GAIA kann beide als zwei getrennte Merkmale einbeziehen: die ursprünglichen Messungen und die Skala 1 bis 9. Sie sind in der Beschreibung der Merkmale assoziiert. Anhand der Kenntnis dieser Assoziation wird, wenn beide vorhanden sind, nur eines beibehalten, um zu vermeiden, daß die Informationen für die Gewichtung zweimal verwendet werden.

### 1.3.6 Beispiel für eine GAIA-Bildschirmkopie

The screenshot shows the GAIA software interface with the following components:

- List of comparisons:** A table with columns: Comparison, Type of comparison, Name of the comparison, Species, and Session. Comparison 1 is selected.
- Display tree:** A hierarchical tree view on the left showing 'Comparison with a threshold of 6', 'Comparison Qualit. + Electr.', 'Distinct varieties [3]', and 'NON-distinct varieties [49]'. Varieties 107, 112, and 113 are highlighted.
- Results of qualitative comparison:** A table with columns: Chara, Long name, Weighting, Note Std/Cycle 1, Note Ref/Cycle 1, Note Std/Cycle 2, and Note Ref/Cycle 2. The table shows characteristics like 'Green color of leaf', 'Number of lobes', 'Time of flowering', 'Length of petals', 'Height', and 'Intensity of yellow color'.

Comparison	Type of comparison	Name of the comparison	Species	Session
1	Qualit. + Electr.	1st year of study	Rapeseed	Threshold 6
5	Qualitative	Qualitative 1st year threshold 12	Rapeseed	Threshold 12
6	Qualit. + Electr. + Quantit.	Variety 84	Rapeseed	Threshold 12

Chara	Long name	Weighting	Note Std/Cycle 1	Note Ref/Cycle 1	Note Std/Cycle 2	Note Ref/Cycle 2
4	Green color of leaf	1,00	5	5	6	5
6	Number of lobes	0,00	5	5	4	5
11	Time of flowering	1,00	5	4	4	4
13	Length of petals	0,00	5	5	4	5
17	Height	0,00	4	5	6	5
82	Intensity of yellow color	0,00	5	6	6	5

1.3.6.1 Der obere Teil *List of comparisons* zeigt drei verschiedene Berechnungen, die in der Datenbank behalten wurden. Vergleich 1 ist markiert (ausgewählt) und auf dem Anzeigebaum angezeigt.

1.3.6.2 Der *Display tree* auf der linken Seite zeigt die Ergebnisse für eine [qualitative + elektrophoretische (bei einer Schwelle von 6)] Berechnung.

1.3.6.3 *Distinct cultivars [3]* zeigt, daß drei Sorten als von allen übrigen unterscheidbar ausgewiesen wurden. Insgesamt befanden sich 52 (49 + 3) Zuchtpflanzen in der Berechnung.

1.3.6.4 Der Anzeigebaum wird benutzt für das Navigieren durch alle möglichen Paare.

1.3.6.5 Der Benutzer kann die Zweige des Baums gemäß seinem Bedarf erweitern oder reduzieren.

1.3.6.6 *NON-distinct cultivars [49]*. Es wurden 49 Zuchtpflanzen mit einem Schwellenwert von 6 für „nicht von allen übrigen unterscheidbar“ befunden.

1.3.6.7 Die erste Sorte, *Variety 107*, hat nur drei verwandte Sorten, während die zweite, *Variety 112*, neun verwandte Sorten, die dritte, *Variety 113*, vier verwandte Sorten usw. hat.

1.3.6.8 *Variety 112* [1][9] gibt an, daß sich die Sorte 112 im ersten Prüfungsjahr befindet [1] und gemäß dem Schwellenwert von 6 neun verwandte Sorten [9] hat.

1.3.6.9 [dist=3.5]*Variety 26* [2] gibt an, daß die Sorte 26 (Vergleich hervorgehoben=ausgewählt) einen GAIA-Abstand von 3,5 von der Sorte 112 aufweist, die sich im zweiten Prüfungsjahr befindet.

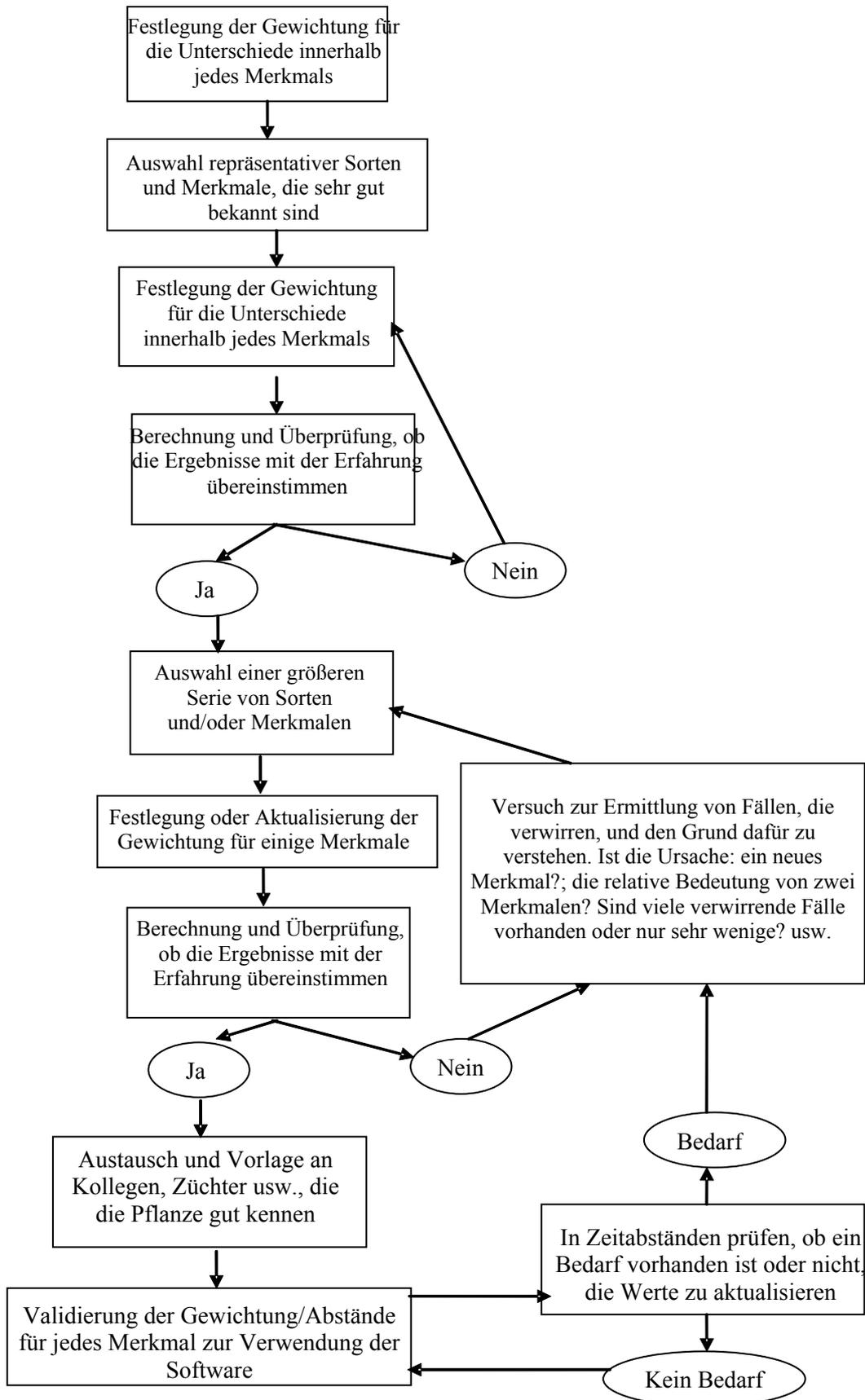
1.3.6.10 Auf der rechten Seite des Anzeigebaums sind die Rohdaten für *Variety 112* und *Variety 26* für die sechs qualitativen Merkmale sichtbar, die an beiden Sorten erfaßt wurden.

1.3.6.11 Die dritte Spalte „Gewichtung“ ist die Gewichtung gemäß den prädefinierten Matrizes. Die Noten für beide Sorten werden für die beiden verfügbaren Perioden gezeigt (Std steht für „*studied*“, was die Kandidatensorten sind).

1.3.6.12 Wie in Rot angezeigt, wird dieses Merkmal nicht angezeigt, wenn zwei Sorten dieselbe Beschreibung für ein Merkmal haben.

1.3.6.13 In dieser Bildschirmkopie wurden die Sorten um der Vertraulichkeit willen numeriert. Die Pflanzensachverständigen können die Sorten gemäß ihrem Bedarf benennen (Parzelle oder Antragsnummer, Name usw.).

Diagramm 3: „Versuchs-und-Prüf“-Prozeß zur Festlegung und Revision der Gewichtung für eine Pflanze



## **2. ELTERNFORMEL VON HYBRIDSORTEN**

### **2.1 Einleitung**

2.1.1 Bei der Prüfung der Unterscheidbarkeit von Hybridsorten können die Behörden die Möglichkeit erwägen, den in diesem Abschnitt beschriebenen Ansatz der Elternformel anzuwenden. Wenn die Ansicht herrscht, daß die Verwendung der Elternformel angebracht sein könnte, wird diese Möglichkeit in den Prüfungsrichtlinien erwähnt.

2.1.2 Die Verwendung der Elternformel setzt voraus, daß der Unterschied zwischen Elternlinien ausreichend ist, um sicherzustellen, daß die aus diesen Eltern erzeugte Hybride unterscheidbar ist. Das Verfahren beruht auf folgenden Schritten:

- i) Beschreibung der Elterlinien gemäß den Prüfungsrichtlinien;
- ii) Prüfung der Eigenständigkeit der Elterlinien im Vergleich zur Sortensammlung auf der Grundlage der Merkmalstabelle in den Prüfungsrichtlinien, um die ähnlichsten Elternlinien zu ermitteln;
- iii) Prüfung der Eigenständigkeit der Hybridformel in bezug auf die Hybriden in der Sortensammlung unter Berücksichtigung der ähnlichsten Elternlinien, und
- iv) Bestimmung der Unterscheidbarkeit an der Hybride bei Sorten mit ähnlicher Formel.

### **2.2 Voraussetzungen des Verfahrens**

Die Anwendung des Verfahrens setzt voraus:

- i) Angabe der Formel und die Einreichung des Vermehrungsmaterials der Elternlinien von Hybridsorten;
- ii) Einbeziehung der als Eltern bei den Hybridsorten der Sortensammlung verwendeten Elternlinien in die Sortensammlung (für Anleitung über die Errichtung einer Sortensammlung vergleiche Dokument TGP/4 Abschnitt 1) und eine Liste der Formeln der Hybridsorten;
- iii) Anwendung des Verfahrens auf alle Sorten in der Sortensammlung. Diese Voraussetzung ist wichtig zur Erzielung des vollen Nutzens, und
- iv) strenges Vorgehen zur Prüfung der Eigenständigkeit einer neuen Elternlinie, um Gewißheit bezüglich der Unterscheidbarkeit der Hybridsorte aufgrund dieser Elternlinie zu erlangen.

### **2.3 Prüfung der Eigenständigkeit einer neuen Elternlinie**

2.3.1 Die Eigenständigkeit einer Elternlinie wird anhand der in den entsprechenden Prüfungsrichtlinien enthaltenen Merkmale geprüft.

2.3.2 Der Unterschied zwischen Elternlinien muß hinreichend sein, um sicherzugehen, daß eine aus ihnen hervorgebrachte Hybride unterscheidbar ist. Beispielsweise:

Merkmal 1: Ein Merkmal mit zwei Ausprägungsstufen (fehlend/vorhanden), die durch zwei Allele eines einzigen Gens bestimmt werden, mit einem dominanten Allel (+) für die Ausprägung „vorhanden“ und einem rezessiven Allel (-) für die Ausprägung „fehlend“.

Drei Elternlinien:

A: mit dem rezessiven Allel (-) für die Ausprägung „fehlend“

B: mit dem dominanten Allel (+) für die Ausprägung „vorhanden“

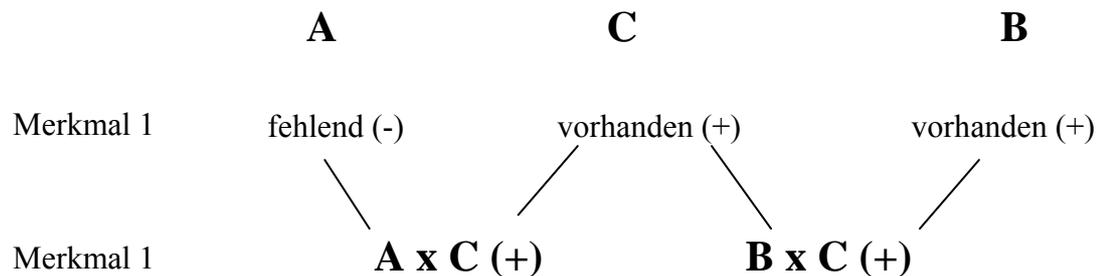
C: mit dem dominanten Allel (+) für die Ausprägung „vorhanden“

Kreuzung der obenerwähnten Elternlinien zur Züchtung folgender F1-Hybriden:

(A x C): mit der Ausprägung „vorhanden“ für Merkmal 1

(B x C): mit der Ausprägung „vorhanden“ für Merkmal 1

Das nachstehende Diagramm zeigt die Art und Weise, wie die zwei verschiedenen Kreuzungen zur gleichen Ausprägung des Merkmals 1 führen (d. h. „vorhanden“ bei beiden Hybriden), obwohl die Elternlinie A(-) und die Elternlinie B(+) unterschiedliche Ausprägungen haben.



2.3.3 Obwohl die Elternlinien A und B für das Merkmal 1 deutlich verschieden sind, haben die beiden Hybridsorten A x C und B x C dieselbe Ausprägung. Somit ist ein Unterschied zwischen A und B für das Merkmal 1 nicht ausreichend.

2.3.4 Mit einer komplexeren genetischen Kontrolle, die mehrere nicht genau beschriebene Gene einbezieht, könnte die Interaktion zwischen den verschiedenen Allelen jedes Gens und zwischen den Genen ebenfalls zu einer ähnlichen Ausprägung an den Hybridsorten führen. In diesen Fällen ist für die Begründung der Unterscheidbarkeit zwischen zwei Elternlinien ein größerer Unterschied angebracht.

2.3.5 Die Bestimmung des erforderlichen Unterschieds beruht hauptsächlich auf einer angemessenen Kenntnis der Art, der Merkmale und, sofern verfügbar, ihrer genetischen Kontrolle.

## **2.4 Überprüfung der Formel**

2.4.1 Zweck der Überprüfung der Formel ist es zu prüfen, ob die hybride Kandidatensorte durch Kreuzung der vom Antragsteller angegebenen und eingereichten Elternlinien hervorgebracht wurde.

2.4.2 Zur Durchführung dieser Überprüfung können verschiedene Merkmale verwendet werden, wenn das genetische Muster jedes Elters in der Hybride identifiziert werden kann. In der Regel können Merkmale verwendet werden, die auf Polymorphismus der Enzyme oder einiger Speicherproteine basieren.

2.4.3 Sind keine geeigneten Merkmale verfügbar, besteht die einzige Möglichkeit darin, die Elternlinien unter Verwendung des vom Antragsteller eingereichten Vermehrungsmaterials zu kreuzen und die Saatgutpartien der Hybridsorte (die vom Antragsteller eingereichte Stichprobe und die nach der Kreuzung geerntete Stichprobe) miteinander zu vergleichen.

## **2.5 Homogenität und Beständigkeit der Elternlinien**

2.5.1 Die Homogenität und die Beständigkeit der Elternlinien sollten gemäß den entsprechenden Empfehlungen für die betreffende Sorte geprüft werden. Die Homogenität und die Beständigkeit der Elternlinien sind wichtig für die Beständigkeit der Hybride. Eine weitere Voraussetzung für die Beständigkeit der Hybride ist die Anwendung derselben Formel für jeden Zyklus der Hybridsaatguterzeugung.

2.5.2 Eine Überprüfung der Homogenität an der Hybride sollte auch dann vorgenommen werden, wenn die Unterscheidbarkeit der Hybride aufgrund der Elternlinien begründet wurde.

## **2.6 Beschreibung der Hybride**

Eine Beschreibung der Hybridsorte sollte auch dann erstellt werden, wenn die Unterscheidbarkeit der Hybride aufgrund der Formel der Elternlinien begründet wurde.

### **3. DAS KOMBINIERTES UNTERSCHIEDBARKEITSKRITERIUM ÜBER MEHRERE JAHRE (COYD)**

#### **3.1 Zusammenfassung der Voraussetzungen für die Anwendung des Verfahrens**

COYD ist ein geeignetes Verfahren für die Prüfung der Unterscheidbarkeit von Sorten, wenn

- das Merkmal quantitativ ist;
- es Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte gibt;
- die Beobachtungen auf Pflanzenbasis (oder Parzellenbasis) über mindestens zwei Jahre erfolgen und an einem einzelnen Prüfungsort durchgeführt werden.
- Es sollte mindestens 20 Freiheitsgrade für das mittlere Abweichungsquadrat Sorten-x-Jahre bei der COYD-Varianzanalyse geben, oder, wenn dies nicht der Fall ist, kann die Langzeit-COYD angewandt werden (vergleiche 3.6.2 unten).

#### **3.2 Zusammenfassung**

3.2.1 Dokument TGP/9/1, Abschnitt 5.2.4.5.1.1 erläutert: „Zur Prüfung der Unterscheidbarkeit von Sorten aufgrund eines quantitativen Merkmals ist es möglich, einen Mindestabstand zwischen Sorten so zu berechnen, daß diese Sorten in bezug auf dieses Merkmal als ‚unterscheidbar‘ angesehen werden können, wenn der berechnete Abstand zwischen einem Sortenpaar größer ist als dieser Mindestabstand. Zu den möglichen Verfahren zur Feststellung der Mindestabstände gehört das kombinierte Unterscheidbarkeitskriterium über mehrere Jahre (COYD)“. Die COYD-Analyse berücksichtigt die Variation zwischen Jahren. Sie ist hauptsächlich für fremdbefruchtende Sorten, einschließlich synthetischer Sorten, bestimmt, kann nach Bedarf unter bestimmten Umständen jedoch auch für selbstbefruchtende und vegetativ vermehrte Sorten verwendet werden. Diese Methode erfordert, daß die Größe der Unterschiede über die Jahre hinreichend stabil ist, und berücksichtigt die Variation zwischen den Jahren.

3.2.2 Das COYD-Verfahren umfaßt:

- für jedes Merkmal die Bildung der Sortenmittelwerte von zwei oder drei Prüfungsjahren für Kandidatensorten und eingetragene Sorten und die Erstellung eines Mittelwertes über die Jahre für die Sorten;
- die Berechnung einer kleinsten gesicherten Differenz (LSD) aufgrund der Variation Sorten-x-Jahre für den Vergleich der Sortenmittelwerte;
- für den Fall, daß die mittlere Differenz über die Jahre zwischen zwei Sorten größer als die LSD ist, die Aussage, daß die Sorten in bezug auf dieses Merkmal als unterscheidbar angesehen werden.

3.2.3 Die hauptsächlichlichen Vorteile des COYD-Verfahrens sind:

- es kombiniert Informationen von mehreren Anbauperioden auf einfache und unkomplizierte Weise in ein einziges Kriterium (das „COYD-Kriterium“);
- es stellt sicher, daß die Beurteilungen der Unterscheidbarkeit in anderen Anbauperioden wiederholbar sind; mit anderen Worten sollte das gleiche genetische Material innerhalb angemessener Grenzen von Anbauperiode zu Anbauperiode ähnliche Ergebnisse bringen;
- das Risiko einer falschen Entscheidung über die Unterscheidbarkeit ist für alle Merkmale konstant.

### **3.3 Einleitung**

Die nachstehenden Abschnitte enthalten Informationen über:

- die dem Verfahren COYD zugrundeliegenden Grundsätze;
- die UPOV-Empfehlungen zur Anwendung von COYD auf einzelne Arten;
- Einzelheiten über die Art und Weise, wie das Verfahren zur Behandlung besonderer Umstände angepaßt werden kann, u. a. wenn sich im Anbauversuch nur eine geringe Anzahl Sorten befindet;
- die für die Anwendung des Verfahrens verfügbare Computersoftware.

### **3.4 Die COYD-Methode**

3.4.1 Die COYD-Methode versucht für jedes Merkmal einen Mindestabstand zu ermitteln, oder einen Abstand der, sofern er von zwei Sorten in Anbauversuchen über eine Zeitspanne von zwei oder drei Jahren erreicht wird, es ermöglicht, diese Sorten mit einem spezifizierten Zuverlässigkeitsgrad als deutlich unterscheidbar anzusehen.

3.4.2 Die Methode verwendet die Variation in der Sortenausprägung eines Merkmals von Jahr zu Jahr, um den Mindestabstand zu ermitteln. So werden Sorten, die eine Konstanz in der Sortenrangfolge zwischen Jahren aufweisen, einen geringeren Mindestabstand als Sorten mit ausgeprägten Änderungen in der Rangfolge aufweisen.

3.4.3 Die Berechnung des COYD-Kriteriums umfaßt eine Varianzanalyse aufgrund der Tabelle der Sorte-x-Jahresmittelwerte für jedes Merkmal, um eine Schätzung der Variation Sorten-x-Jahre zu erzielen, die im nächsten Schritt verwendet wird: Berechnung einer LSD. Die Daten für alle Kandidatensorten und etablierten Sorten, die sich in den Anbauversuchen über die zwei oder drei Prüfungsjahre ergeben, sind in der Regel in der Tabelle eingeschlossen, die Analyse erfolgt mittels einer Varianzanalyse, die Standardabweichung wird als Schätzung der Variation Sorten-x-Jahre verwendet, und die sich daraus ergebende LSD wird als COYD-LSD bezeichnet. Wenn die Zahl der Sorten im Anbauversuch gering ist, wird jedoch ein anderes Vorgehen befolgt.

3.4.4 Ist die Zahl der Sorten im Anbauversuch gering, wird die zur Berechnung des COYD-Kriteriums angewandte Tabelle um die Mittelwerte anderer Sorten und früherer Jahre erweitert, ein anderes Analyseverfahren angewandt, um eine Standardabweichung Sorten-x-Jahre für die Schätzung der Variation Sorten-x-Jahre zu erzielen, und die sich daraus ergebende LSD wird als Langzeit-LSD bezeichnet. Diese werden später behandelt.

3.4.5 Gleichung [1]

$$\text{LSD}_p = t_p \times \sqrt{2} \times \boxed{\phantom{000}}$$

wobei  $\boxed{\phantom{000}}$  der Standardfehler eines Sortenmittelwerts über ein Jahr ist, berechnet als:

$$\boxed{\phantom{000}} =$$

$$\sqrt{\frac{\text{Sorten mal Mittelquadrat der Jahre}}{\text{Anzahl der Prüfungsjahre}}}$$

und  $t_p$  der Wert in der Student-t-Tabelle ist, geeignet für eine zweiseitige Prüfung mit Wahrscheinlichkeit  $p$  und mit dem mittleren Abweichungsquadrat der Sorten-x-Jahre assoziierten Freiheitsgraden. Auf das Wahrscheinlichkeitsniveau  $p$ , das für einzelne Arten angemessen ist, wird unter UPOV-EMPFEHLUNGEN ZU COYD unten eingegangen.

3.4.6 Ein Beispiel für die Anwendung von COYD für eine kleine Menge von Datensätzen ist in Abbildung 1 wiedergegeben. Statistische Einzelheiten des Verfahrens sind in Teil II Abschnitt 3.9 zu entnehmen. Für weitere Informationen über das COYD-Kriterium wird auf eine Veröffentlichung von Patterson und Weatherup (1984) verwiesen.

### **3.5 Anwendung von COYD**

3.5.1 COYD ist ein geeignetes Verfahren für die Prüfung der Unterscheidbarkeit von Sorten, wenn

- das Merkmal quantitativ ist;
- es Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte gibt;
- die Beobachtungen auf Pflanzenbasis (oder Parzellebasis) über zwei oder mehrere Jahre erfolgen.
- Es sollte mindestens 20 Freiheitsgrade für das mittlere Abweichungsquadrat Sorten-x-Jahre bei der COYD-Varianzanalyse geben, oder, wenn dies nicht der Fall ist, kann die Langzeit-COYD angewandt werden (vergleiche 3.6.2 unten).

3.5.2 Ein Sortenpaar gilt als unterscheidbar, wenn sich seine Mittelwerte über mehrere Jahre in einem oder mehreren Merkmalen zumindest in der COYD-LSD unterscheiden.

3.5.3 Das von der UPOV empfohlene Wahrscheinlichkeitsniveau  $p$  für den Wert  $t_p$ , der für die Berechnung der COYD LSD benutzt wird, ist je nach Pflanze unterschiedlich und hängt für einige Pflanzen davon ab, ob die Prüfung über zwei oder drei Jahre durchgeführt wird. Die Prüfungsschemata, die sich in der Regel bei der Prüfung der Unterscheidbarkeit ergeben, sind in Dokument TGP/8/1 Teil II Abschnitt 3.11 beschrieben.

### **3.6 Anpassung von COYD an besondere Gegebenheiten**

#### **3.6.1 Unterschiede zwischen Jahren im Ausprägungsbereich eines Merkmals**

Gelegentlich können zwischen Jahren deutliche Unterschiede im Ausprägungsbereich eines Merkmals auftreten. So kann sich z. B. in einem späten Frühling die Zeitspanne des Ährenschiebens von Gräsern verkürzen. Um diesem Effekt Rechnung zu tragen, können bei der Varianzanalyse besondere Größen – eine für jedes Jahr – vorgesehen werden. Jede Größe stellt die lineare Regression der Beobachtungen für das betreffende Jahr gegenüber dem Sortenmittelwert über alle Jahre dar. Dieses Verfahren ist als die modifizierte kombinierte Regressionsanalyse (Modified Joint Regression Analyse) (MJRA) bekannt und wird in Situationen empfohlen, in denen ein statistisch signifikanter ( $p \leq 1\%$ ) Beitrag der Regressionsgrößen in der Varianzanalyse vorhanden ist. Statistische Einzelheiten und ein Computerprogramm zur Ausführung des Verfahrens sind in Teil II Abschnitte 3.9 und 3.10 zu finden.

#### **3.6.2 Geringe Anzahl von Sorten in den Prüfungen: Langzeit-COYD**

3.6.2.1 Es wird empfohlen, bei der COYD-Varianzanalyse mindestens 20 Freiheitsgrade für das mittlere Abweichungsquadrat Sorten-x-Jahre zu haben, um sicherzustellen, daß das mittlere Abweichungsquadrat der Sorten-x-Jahre auf ausreichenden Daten beruht, um eine zuverlässige Schätzung der Variation Sorten-x-Jahre für die LSD zu ergeben. Zwanzig Freiheitsgrade entsprechen 11 Sorten, die üblicherweise in Dreijahresprüfungen geprüft werden, oder 21 Sorten in Zweijahresprüfungen). Anbauversuche mit weniger gemeinsam vorhandenen Sorten über die Jahre werden so angesehen, daß sie eine geringe Anzahl Sorten prüfen.

3.6.2.2 In diesen Anbauversuchen können die Mittelwerttabellen Sorte-x-Jahr erweitert werden, um die Mittelwerte früherer Jahre und nach Bedarf anderer etablierter Sorten einzubeziehen. Da nicht alle Sorten in allen Jahren vorhanden sind, sind die sich ergebenden Mittelwerttabellen Sorte-x-Jahr nicht symmetrisch. Infolgedessen wird jede Tabelle mit dem Verfahren der kleinsten Quadrate von Ausgleichskonstanten (FITCON) oder mit REML analysiert, die eine alternative Standardabweichung Sorten-x-Jahre als langfristige Schätzung der Variation Sorte-x-Jahre ergibt. Diese Schätzung weist mehr Freiheitsgrade auf, da sie auf mehr Jahren und Sorten beruht.

$$\text{Freiheitsgrade} = \left( \frac{\text{Anzahl Werte in der erweiterten}}{\text{Tabelle Sorten - x - Jahre}} \right) - (\text{Anzahl Sorten}) - (\text{Anzahl Jahre}) + 1$$

3.6.2.3 Das alternative mittlere Abweichungsquadrat Sorten-x-Jahre wird in der obigen Gleichung [1] für die Berechnung einer LSD benutzt. Diese LSD wird als „Langzeit-LSD“ bezeichnet, um sie von der COYD-LSD zu unterscheiden, die lediglich auf Prüfungsjahren und Sorten beruht. Die Langzeit-LSD wird auf dieselbe Weise angewandt wie die COYD-LSD für die Unterscheidbarkeitsprüfung von Sorten durch Vergleich ihrer Mittelwerte über die Jahre (Prüfungsjahre). Das Verfahren des Vergleichs der Mittelwerte der Sorten anhand der „Langzeit-LSD“ wird als „Langzeit-COYD“ bezeichnet.

3.6.2.4 Die Langzeit-COYD sollte nur auf diejenigen Merkmale angewandt werden, die nicht die empfohlenen Mindestfreiheitsgrade aufweisen. Liegt jedoch ein Nachweis dafür vor, daß die LSD eines Merkmals über die Jahre deutlich schwankt, kann es notwendig sein, die LSD für dieses Merkmal auf die aktuellen Daten der zwei oder drei Jahre zu stützen, selbst wenn es wenige Freiheitsgrade aufweist.

3.6.2.5 Abbildung 2 gibt ein Beispiel für die Anwendung der Langzeit-COYD auf Italienisches Weidelgras am Merkmal „Wuchsform im Frühjahr“. Ein Datenflußdiagramm der Stadien und DUST-Module, das zur Erzeugung der Langzeit-LSD und Ausführung der Langzeit-COYD verwendet wird, ist in Abbildung B2 in Teil II: Abschnitt 3.10 wiedergegeben.

### 3.6.3 Deutliche Änderungen von Jahr zu Jahr bei einem einzelnen Sortenmerkmal

Gelegentlich kann ein Sortenpaar auf der Grundlage eines t-Tests, der allein wegen des sehr großen Unterschieds zwischen den Sorten in einem einzelnen Jahr signifikant ist, für unterscheidbar erklärt werden. Um diese Situationen zu überwachen, wird eine Kontrollkenngröße, genannt  $F_3$ , berechnet – d. h. das mittlere Abweichungsquadrat Sorte-x-Jahre für das betreffende Sortenpaar, ausgedrückt als Quotient des allgemeinen mittleren Abweichungsquadrats der Sorte-x-Jahre. Diese Kenngröße sollte mit F-Verteilungstabellen mit 1 und  $g$  bzw. 2 und  $g$  Freiheitsgraden für Prüfungen mit Daten für zwei bzw. drei Jahre verglichen werden, wobei  $g$  die Freiheitsgrade des mittleren Abweichungsquadrates Sorte-x-Jahre darstellt. Übersteigt der berechnete  $F_3$ -Wert den tabellierten F-Wert auf dem 1%-Niveau, dann sollte eine Erklärung für das ungewöhnliche Ergebnis gesucht werden, bevor eine Entscheidung über die Unterscheidbarkeit getroffen wird.

## 3.7 Umsetzung von COYD

COYD ist ein geeignetes Verfahren für die Prüfung der Unterscheidbarkeit von Sorten, wenn

- das Merkmal quantitativ ist;
- es Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte gibt;
- die Beobachtungen auf Pflanzenbasis (oder Parzellenbasis) über zwei oder mehrere Jahre erfolgen.

- Es sollte mindestens 20 Freiheitsgrade für das Quadrat des Mittelwertes Sorten-x-Jahre bei der COYD-Varianzanalyse geben, oder, wenn dies nicht der Fall ist, kann die Langzeit-COYD angewandt werden (vergleiche 3.6.2 oben).

Das COYD-Verfahren kann anhand des TVRP-Moduls des DUST-Pakets für die statistische Analyse von DUS-Daten angewandt werden, das von Frau Sally Watson (Email: [info@afbini.gov.uk](mailto:info@afbini.gov.uk)) oder unter <http://www.afbini.gov.uk/dustnt.htm> bezogen werden kann. Muster der Ausgabedaten sind in Teil II Abschnitt 3.10 angegeben.

### **3.8 Quellenangaben**

DIGBY, P.G.N. (1979). Modified joint regression analysis for incomplete variety x environment data. J. Agric. Sci. Camb. 93, 81-86.

PATTERSON, H.D. & WEATHERUP, S.T.C. (1984). Statistical criteria for distinctness between varieties of herbage crops. J. Agric. Sci. Camb. 102, 59-68.

TALBOT, M. (1990). Statistical aspects of minimum distances between varieties. UPOV-TWC-Dokument TWC/VIII/9, UPOV, Genf.

**Abbildung 1: Veranschaulichung der Anwendung des COYD-Kriteriums**

**Merkmal: Tage bis zum Ährenschieben bei Sorten von Deutschem Weidelgras**

Sorten	Jahre			Mittelwert über die Jahre	Unterschied (Sorten verglichen mit C2)	
	1	2	3			
<i>Vergleichssorte</i>		Mittelwerte				
R1	38	41	35	38	35	D
R2	63	68	61	64	9	D
R3	69	71	64	68	5	D
R4	71	75	67	71	2	
R5	69	78	69	72	1	
R6	74	77	71	74	-1	
R7	76	79	70	75	-2	
R8	75	80	73	76	-3	
R9	78	81	75	78	-5	D
R10	79	80	75	78	-5	D
R11	76	85	79	80	-7	D
<i>Kandidatensorte</i>						
C1	52	56	48	52	21	D
C2	72	79	68	73	0	-
C3	85	88	85	86	-13	D

**VARIANZANALYSE**

Ursache	df	mittleres Abweichungsquadrat
Jahre	2	174,93
Sorte	13	452,59
Sorte-x-Jahre	26	2,54

$$LSD_p = t_p * \sqrt{2} * SE(\bar{X})$$

$$LSD_{0,01} = 2,779 * 1,414 * \sqrt{(2,54/3)} = 3,6$$

wobei  $t_p$  der Student-t-Tabelle entnommen ist mit  $p = 0,01$  (zweiseitig) und 26 Freiheitsgraden

Um die Unterscheidbarkeit einer Kandidatensorte zu bestimmen, wird der Unterschied der Mittelwerte zwischen ihr und allen anderen Sorten berechnet. In der Praxis wird für jede Kandidatensorte eine Spalte der Unterschiede berechnet. In diesem Fall werden Sorten mit Mittelwertunterschieden größer als oder gleich wie 3,6 als deutlich unterscheidbar (oben mit D markiert) angesehen.

**Abbildung 2: Veranschaulichung der Anwendung der Langzeit-COYD**

**Merkmal: Wuchsform im Frühjahr bei Sorten von Italienischem Weidelgras**

Sorten	1	2	Jahre			Mittelwert über die Prüfungs- jahre	<i>Unterschied (Sorten verglichen mit C2)</i>		
			3*	4*	5*				
<i>Vergleichssorte</i>	Mittelwerte								
R1	43	42	41	44					
R2		39	45						
R3	43	38	41	45	40	42	6	D	
R4	44	40	42	48	44	44,7	3,3	D	
R5	46	43	48	49	45	47,3	0,7		
R6	51	48	52	53	51	52	-4	D	
<i>Kandidatensorte</i>									
C1			43	45	44	44	4	D	
C2			49	50	45	48	0		
C3			48	53	47	49,3	-1,3		

\* gibt ein Prüfungsjahr an

Ziel ist die Prüfung der Unterscheidbarkeit der Kandidatensorten C1, C2 und C3, die in den Prüfungsjahren 3, 4 und 5 angebaut werden.

Der Anbauversuch zählt eine geringe Anzahl zu prüfende Sorten, weil lediglich sieben Sorten über die Prüfungsjahre 3, 4 und 5 gemeinsam vorhanden sind (die Daten sind mit schwarzem Rand markiert).

Die FITCON-Analyse der auf neun Sorten in fünf Jahren erweiterten Mittelwerttabelle Sorte-x-Jahre ergibt: mittleres Abweichungsquadrat Sorte-x-Jahre = 1,924 bei 22 Freiheitsgraden.

$$\text{Langzeit-LSD}_p = t_p * \sqrt{2} * \text{SE}(\bar{X})$$

$$\text{Langzeit-LSD}_{0,01} = 2,819 * 1,414 * \sqrt{(1,924/3)} = 3,19$$

wobei  $t_p$  der Student-t-Tabelle mit  $p = 0,01$  (zweiseitig) und 22 Freiheitsgraden entnommen wird.

Zur Prüfung der Unterscheidbarkeit einer Kandidatensorte wird der Unterschied bei den Mittelwerten zwischen der Kandidatensorte und allen übrigen Sorten berechnet. In der Praxis wird eine Spalte mit Unterschieden für jede Kandidatensorte berechnet. Im Falle der Sorte C2 werden Sorten mit durchschnittlichen Unterschieden, die größer als oder gleich 3,19 sind, als unterscheidbar angesehen (oben mit D bezeichnet).

### **3.9 Statistische Verfahren der COYD**

#### **3.9.1 Varianzanalyse**

Die beim COYD-Kriterium verwendeten Standardfehler basieren auf der Varianzanalyse der Sorten-x-Jahre-Tabelle der Merkmalsmittelwerte. Für  $m$  Jahre und  $n$  Sorten teilt die Varianzanalyse die verfügbaren Freiheitsgrade wie folgt auf:

Ursache	df
Jahre	$m-1$
Sorten	$n-1$
Sorten-x-Jahre	$(m-1)(n-1)$

#### **3.9.2 Modifizierte kombinierte Regressionsanalyse (MJRA)**

3.9.2.1 Wie oben erwähnt, verwendet das COYD-Kriterium die anhand des mittleren Abweichungsquadrats Sorten-x-Jahre geschätzte Sorten-x-Jahre-Variation als Basis für den Standardfehler eines Sortenmittelwerts. Eine systematische Variation kann bisweilen auch als nichtsystematische Variation ausgewiesen werden. Dieser systematische Effekt verursacht das Auftreten verschiedener Anstiegswinkel der Regressionskurve der Sortenmittelwerte in einzelnen Jahren im Vergleich zu den durchschnittlichen Sortenmittelwerten über alle Jahre. Dieser Effekt kann für das Merkmal „Zeitpunkt des Ährenschiebens“ in einem Jahr mit spätem Frühling festgestellt werden: Die Spannweite der Zeitpunkte des Ährenschiebens kann im Vergleich zur normalen komprimiert sein. Dies führt zu einer Reduzierung des Anstiegs der Regressionskurve im Verhältnis zum Durchschnitt der Sortenmittelwerte. Eine nichtsystematische Variation wird durch die Variation über diese Regressionskurven dargestellt. Wenn nur nichtsystematische Sorten-x-Jahresvariationen auftreten, haben die Anstiege der Regressionskurven in allen Jahren den konstanten Wert 1,0. Wenn jedoch systematische Variationen vorhanden sind, treten Anstiege auf, die von 1,0 abweichen, jedoch mit einem Durchschnittswert von 1,0. Bei der Anwendung von MJRA beruht der Standardfehler eines Sortenmittelwerts auf dem nichtsystematischen Teil der Sorten-x-Jahr-Variation.

3.9.2.2 Die Unterscheidung zwischen der gesamten Sorten-x-Jahresvariation und der durch MJRA bereinigten Sorten-x-Jahresvariation ist in Abbildung B1 wiedergegeben, in der die Sortenmittelwerte in jedem der drei Jahre mit dem Durchschnitt der Sortenmittelwerte aller Jahre verglichen werden. Die Variation von drei an die Daten angepaßten parallelen Kurven, eine für jedes Jahr, ergibt die Gesamtvariation Sorten-x-Jahre, wie in den oben beschriebenen COYD-Kriterien verwendet. Diese Regressionskurven haben den gemeinsamen Anstieg 1,0. Diese Variation kann durch Bereinigung zusätzlicher getrennter Regressionskurven für jedes Jahr reduziert werden. Die resultierende Restvariation über die individuellen Regressionskurven liefert das MJRA-bereinigte mittlere Abweichungsquadrat Sorten-x-Jahre, auf dem der Standardfehler für ein Sortenmittel beruhen könnte. Es ist sichtbar, daß diese die MJRA-Bereinigung nur wirksam ist, wenn sich die Anstiege der Regressionskurven zwischen den Jahren unterscheiden, wie dies bei den Zeitpunkten des Ährenschiebens der Fall sein kann.

3.9.2.3 Die Anwendung dieses Verfahrens bei der Prüfung der Unterscheidbarkeit wurde als Option in das Computerprogramm aufgenommen, das das COYD-Kriterium im DUST-Paket anwendet. Es wird empfohlen, es nur anzuwenden, wenn sich die Anstiege der Sortenregressionskurven zwischen den Jahren mit 1 % Signifikanzniveau signifikant unterscheiden. Dieses Niveau kann im Computerprogramm angegeben werden.

3.9.2.4 Zur Berechnung der bereinigten Sortenmittelwerte und der Anstiege der Regressionskurven wird von folgendem Modell ausgegangen:

$$y_{ij} = u_j + b_j v_i + e_{ij}$$

wobei  $y_{ij}$  den Wert für die  $i$ . Sorte im  $j$ . Jahr darstellt.

$u_j$  ist der Mittelwert des Jahres  $j$  ( $j = 1, \dots, m$ )

$b_j$  ist der Anstieg der Regression für das Jahr  $j$

$v_i$  ist die Wirkung der Sorte  $i$  ( $i = 1, \dots, n$ )

$e_{ij}$  ist ein Fehlerwert.

3.9.2.5 Aus den Gleichungen (6) und (7) von Digby (1979), bei einer Umkehrung der Bedeutung der Jahre und Sorten, erhält man für Situationen mit vollständigen Daten die folgenden Gleichungen hinsichtlich der Werte:

$$\sum_{i=1}^n v_i y_{ij} = b_j \sum_{i=1}^n v_i^2$$

$$\sum_{j=1}^m b_j y_{ij} = v_i \sum_{j=1}^m b_j^2$$

3.9.2.6 Diese Gleichungen werden schrittweise gelöst, indem als Ausgangspunkt alle  $b_j$ -Werte als 1,0 angesetzt werden, um Werte für die  $v_i$  zu erhalten. Die MJRA-Restquadratsumme erhält man aus:

$$\sum_{j=1}^m \sum_{i=1}^n (y_{ij} - u_j - b_j v_i)^2$$

3.9.2.7 Diese Quadratsumme wird zur Berechnung des MJRA-bereinigten mittleren Abweichungsquadrats Sorten-x-Jahre bei  $(m-1)(n-1) - m + 1$  Freiheitsgraden angewandt.

### 3.9.3 Vergleich von COYD mit anderen Kriterien

Es kann gezeigt werden, daß für einen Test über drei Jahre das auf dem Wahrscheinlichkeitsniveau von 1 % angewendete COYD-Kriterium etwa die gleiche Strenge aufweist wie das 2x1 %-Kriterium für ein Merkmal, bei dem die Quadratwurzel des Verhältnisses des mittleren Abweichungsquarates Sorte-x-Jahre zum mittleren

Abweichungsquadrat Sorte-x-Wiederholungen innerhalb der Anbauversuche ( $\lambda$ ) einen Wert von 1,7 hat. Das auf dem 1 %-Niveau angewandte COYD-Kriterium ist weniger streng als das 2x1 %-Kriterium, wenn  $\lambda < 1,7$  ist, und strenger, wenn  $\lambda > 1,7$  ist.

### **3.10 Die COYD-Software**

3.10.1 Ein Beispiel der Ausgabedaten aus dem Computerprogramm im DUST-Paket, das das COYD-Kriterium anwendet, ist in den Tabellen B1 bis B 3 wiedergegeben. Es wurde einem Anbauversuch mit Deutschem Weidelgras (diploid) mit 40 aus der Sortensammlung ausgewählten Sorten (R1 bis R40) und neun Kandidatensorten (C1 bis C9) in sechs Wiederholungen entnommen, an denen acht Merkmale über die Jahre 1988, 1989 und 1990 gemessen wurden.

3.10.2 Jedes der acht Merkmale wird anhand der Varianzanalyse analysiert. Da diese Analyse die Daten Sorte-x-Jahr-x-Wiederholung betrifft, betragen die mittleren Abweichungsquadrate 6 (= Anzahl Wiederholungen) Mal die Größe der mittleren Abweichungsquadrate der Varianzanalyse der Daten Sorte-x-Jahr, auf die im Hauptteil dieses Dokuments hingewiesen wird. Die Ergebnisse sind in Tabelle B1 wiedergegeben. Neben den Sortenmittelwerten-x-Jahre sind ebenfalls angegeben:

YEAR MS:	Größe des mittleren Abweichungsquadrates der Jahre
VARIETY MS:	Größe des mittleren Abweichungsquadrates der Sorten
VAR.YEAR MS:	mittleres Abweichungsquadrat der Sorten-x-Jahre-Interaktion
F1 RATIO:	das Verhältnis VARIETY MS zu VAR.YEAR MS (ein Maß der Unterscheidungskraft des Merkmals – hohe Werte geben eine hohe Unterscheidungskraft an)
VAR.REP MS:	Durchschnitt des mittleren Abweichungsquadrates Sorten-x-Wiederholung aus jedem Jahr
LAMBDA VALUE ( $\lambda$ ):	Quadratwurzel des Verhältnisses VAR. YEAR MS zu VAR. REP MS;
BETWEEN SE:	Standardfehler der Sortenmittelwerte in Anbauversuchen auf Parzellenbasis, d. h. Quadratwurzel aus VAR.YEAR MS, dividiert durch 18 (3 Jahre x 6 Wiederholungen)
WITHIN SE:	Standardfehler der Sortenmittelwerte in einem Anbauversuch auf Parzellenbasis, d. h. Quadratwurzel aus VAR.REP MS, dividiert durch 18
DF:	Freiheitsgrade für Sorten-x-Jahreswerte
MJRA SLOPE:	Steigung der Regression eines Sortenmittelwerts aus einem einzelnen Jahr auf die Mittelwerte über die drei Jahre
REGR F VALUE:	mittleres Abweichungsquadrat der MJRA Regression als Verhältnis des mittleren Abweichungsquadrats um die Regression
REGR PROB:	statistische Signifikanz des REGR F VALUE;
TEST:	gibt an ob MJRA-Bereinigung angewandt (REG) oder nicht angewandt (COY) wurde.

3.10.3 Jede Kandidatensorte wird mit jeder anderen Kandidatensorte und jeder anderen aus der Sortensammlung ausgewählten Sorte im Anbauversuch verglichen. Die mittleren Unterschiede zwischen Sortenpaaren werden mit der LSD für das Merkmal verglichen. Die

Ergebnisse für das Sortenpaar R1 und C1 sind in Tabelle B2 wiedergegeben. Die einzelnen t-Werte innerhalb eines Jahres sind aufgelistet, um Informationen über verschiedene Jahre zu erteilen. Die Sorten R1 und C1 werden als unterscheidbar angesehen, weil zumindest bei einem Merkmal ein mittlerer Unterschied auf dem 1 %-Niveau COYD-signifikant ist. Wenn der  $F_3$ -Quotient für das Merkmal 8 vielmehr auf dem Niveau 1 % als auf dem Niveau 5 % signifikant gewesen wäre, wären die Daten für das Merkmal 8 untersucht worden, und weil die Unterschiede in den drei Jahren nicht alle in dieselbe Richtung weisen, hätte die COYD-Signifikanz für das Merkmal 8 nicht für die Unterscheidbarkeit gezählt.

3.10.4 Das Ergebnis in bezug auf die Prüfung der Unterscheidbarkeit jeder Kandidatensorte von allen anderen Sorten ist in Tabelle B3, wiedergegeben, wo D „unterscheidbar“ und ND „nicht unterscheidbar“ bedeuten.

TGP/8/1 Draft 16: TEIL II: 3: KOMBINIERTES  
UNTERSCHIEDBARKEITSKRITERIUM ÜBER MEHRERE JAHRE

Seite 75

**Tabelle B 1: Ein Beispiel eines Teils der Ausgabedaten des COYD-Programms, das die Sortenmittelwerte und die Varianzanalyse der Merkmale zeigt**

DEUTSCHES WEIDELGRAS (DIPLOID) FRÜH N.I. UPOV 1988-90

	SORTENMITTELWERTE ÜBER DIE JAHRE							
	5	60	8	10	11	14	15	24
	SP.HT	NSPHT	DEEE	H.EE	WEE	LFL	WFL	LEAR
1 R1	45,27	34,60	67,87	45,20	70,05	20,39	6,85	24,54
2 R2	42,63	31,84	73,85	41,96	74,98	19,68	6,67	24,44
3 R3	41,57	27,40	38,47	27,14	57,60	17,12	6,85	22,57
4 R4	33,35	21,80	77,78	30,77	78,04	18,25	6,40	21,09
5 R5	37,81	25,86	50,14	27,24	62,64	16,41	6,41	16,97
6 R6	33,90	21,07	78,73	32,84	79,15	19,44	6,46	21,79
7 R7	41,30	31,37	73,19	41,35	71,87	20,98	6,92	24,31
8 R8	24,48	19,94	74,83	32,10	62,38	15,22	6,36	19,46
9 R9	46,68	36,69	63,99	44,84	68,62	18,11	7,02	22,58
10 R10	25,60	20,96	75,64	32,31	57,20	14,68	5,51	20,13
11 R11	41,70	30,31	74,60	40,17	76,15	19,45	6,79	22,72
12 R12	28,95	21,56	66,12	27,96	59,56	14,83	5,53	20,55
13 R13	40,67	29,47	70,63	36,81	74,12	19,97	7,04	24,05
14 R14	26,68	20,53	75,84	34,14	63,29	15,21	6,37	20,37
15 R15	26,78	20,18	75,54	30,39	66,41	16,34	6,01	20,94
16 R16	42,44	27,01	59,03	30,39	72,71	17,29	6,47	22,48
17 R17	27,94	21,58	76,13	32,53	68,37	16,72	6,11	22,03
18 R18	41,34	30,85	69,80	37,28	69,52	20,68	7,09	25,40
19 R19	33,54	23,43	73,65	30,35	75,54	18,97	6,37	22,43
20 R20	44,14	34,48	68,74	42,60	64,17	18,63	6,56	22,02
21 R21	27,77	21,53	80,52	31,59	69,41	16,81	5,81	22,35
22 R22	38,90	27,83	75,68	43,25	75,08	19,63	7,46	23,99
23 R23	42,43	31,80	72,40	42,07	74,77	20,99	6,78	23,57
24 R24	38,50	27,73	73,19	37,12	75,76	19,28	6,91	22,77
25 R25	43,84	29,60	68,82	39,79	74,83	20,63	7,08	22,65
26 R26	49,48	36,53	63,45	42,01	70,46	22,14	7,84	25,91
27 R27	25,61	19,25	78,78	29,81	56,81	15,81	5,07	18,94
28 R28	26,70	20,31	79,41	32,75	66,54	16,92	6,00	21,91
29 R29	27,90	20,94	72,66	29,85	67,14	16,85	6,28	21,79
30 R30	43,07	30,34	70,53	40,51	73,23	19,49	7,28	23,70
31 R31	38,18	25,47	74,23	36,88	80,23	20,40	7,09	25,21
32 R32	35,15	27,56	71,49	37,26	63,10	18,18	6,80	23,13
33 R33	42,71	31,09	67,58	39,14	70,36	19,85	7,12	23,35
34 R34	23,14	18,05	72,09	24,29	59,37	13,98	5,63	18,91
35 R35	32,75	25,41	77,22	38,90	67,07	17,16	6,42	21,49
36 R36	41,71	31,94	77,98	44,33	73,00	19,72	7,09	23,45
37 R37	44,06	32,99	74,38	45,77	71,59	20,88	7,40	24,06
38 R38	42,65	32,97	74,76	44,42	74,13	20,29	7,38	24,32
39 R39	28,79	22,41	76,83	35,91	64,52	16,85	6,34	22,24
40 R40	44,31	31,38	72,24	43,83	74,73	21,53	7,60	25,46
41 C1	42,42	31,68	64,03	40,22	67,02	20,73	6,90	26,16
42 C2	41,77	32,35	86,11	46,03	75,35	20,40	6,96	22,99
43 C3	41,94	31,09	82,04	43,17	74,04	19,06	6,26	23,44
44 C4	39,03	28,71	78,63	45,97	70,49	21,27	6,67	23,37
45 C5	43,97	30,95	72,99	39,14	77,89	19,88	6,68	25,44
46 C6	37,56	27,14	83,29	39,16	81,18	19,47	6,97	25,25
47 C7	38,41	28,58	83,90	42,53	76,44	19,28	6,00	23,47
48 C8	40,08	27,25	83,50	43,33	80,16	22,77	7,92	26,81
49 C9	46,77	34,87	51,89	37,68	61,16	19,25	6,92	24,82
YEAR MS	1279,09	3398,82	3026,80	2278,15	8449,20	672,15	3,36	51,32
SORTE MS	909,21	476,72	1376,10	635,27	762,41	80,21	6,44	74,17
VAR.YEAR MS	23,16	18,86	14,12	23,16	46,58	4,76	0,28	2,73
F1 RATIO	39,26	25,27	97,43	27,43	16,37	16,84	22,83	27,16
VAR.REP MS	8,83	8,19	4,59	11,95	23,23	1,52	0,15	1,70
LAMBDA VALUE	1,62	1,52	1,75	1,39	1,42	1,77	1,37	1,27
BETWEEN SE	1,13	1,02	0,89	1,13	1,61	0,51	0,13	0,39
WITHIN SE	0,70	0,67	0,50	0,81	1,14	0,29	0,09	0,31
DF	96	94	96	96	96	96	96	96
MJRA SLOPE 88	0,90	0,86	0,99	0,91	0,99	1,09	0,97	0,95
MJRA SLOPE 89	1,05	1,08	1,01	0,99	1,06	0,97	1,02	0,98
MJRA SLOPE 90	1,05	1,06	1,00	1,10	0,95	0,94	1,01	1,07
REGR F VAL	4,66	6,17	0,06	4,48	0,76	1,62	0,29	1,91
REGR PROB	1,17	0,30	93,82	1,39	47,08	20,27	74,68	15,38
TEST	COY	REG	COY	COY	COY	COY	COY	COY

**Tabelle B 2: Ein Beispiel der Ausgabedaten aus dem COYD-Programm, das einen Vergleich der Sorten R1 und C1 zeigt**

DEUTSCHES WEIDELGRAS (DIPLOID) FRÜH N.I. UPOV 1988-90

41 C1 KONTRA 1 R1  
 \*\*\*

\*\*\* MIT REGRESS.-BEREINIGUNG, WENN SIGNIFIKANT

(T-WERTE + VE IF 41 C1 > 1 R1)

	SIG.-NIVEAUS				COYD			T-WERTE				F <sub>3</sub>
	JAHRE				T	WAHR- SCH. %	SIG.	JAHRE			T-BEWERTUNG	
	88	89	90					88	89	90		
5 SP.HGHT	-	-	-1	ND	-1,78	7,88	NS	-1,05	-1,34	-2,64	-2,64	0,23 NS
60 NATSPHT	-	-1	-	ND	-2,02	4,61	*	-1,58	-2,61	-1,17	-2,61	0,22 NS
8 DATEEE	-1	-1	+	D	-3,06	0,29	**	-4,14	-6,33	0,80	-6,74	3,99 *
10 HGHT.EE	-1	-1	-5	D	-3,11	0,25	**	-2,79	-2,69	-2,06	-7,55	0,06 NS
11 WIDTHEE	-	-	-	ND	-1,33	18,58	NS	-1,47	-1,80	-0,21	0,00	0,32 NS
14 LGTHFL	+	+	-	ND	0,47	63,61	NS	0,17	1,83	-0,67	0,00	0,56 NS
15 WIDTHFL	+	-	+	ND	0,27	78,83	NS	0,31	-0,41	0,67	0,00	0,17 NS
24 EARLGTH	5	1	+	ND	2,93	0,42	**	2,10	3,33	1,01	5,43	0,84 NS

### Anmerkungen

- Die drei COYD-Spalten unter den Titeln T WAHR SCH. % SIG geben den COYD-t-Wert, seine Signifikanzwahrscheinlichkeit und sein Signifikanzniveau wieder. Der t-Wert ist die Kenngröße, die sich durch Division der Mittelwertdifferenz zwischen zwei Sorten durch den Standardfehler dieser Differenz ergibt. Der t-Wert kann auf Signifikanz überprüft werden, indem er mit geeigneten Werten der Student-t-Tabelle verglichen wird. Die Berechnung und Prüfung eines t-Wertes auf diese Weise entspricht der Ableitung einer LSD und der Überprüfung, ob die Mittelwertdifferenz größer als die LSD ist.
- Die beiden F<sub>3</sub>-Spalten rechts geben die Kenngröße für den F<sub>3</sub>-Varianzquotienten und sein Signifikanzniveau wieder. Die Kenngröße F<sub>3</sub> wird in Teil II Abschnitt 3.6.2 definiert.
- Die Sektionen in den Kästen beziehen sich auf frühere Unterscheidbarkeitskriterien. Die drei Spalten „T-WERTE, JAHRE“ mit den Überschriften 88, 89 und 90 sind die einzelnen t-Test-Werte innerhalb eines Jahres (der zweiseitige Student-t-Test der Sortenmittelwerte mit Standardfehlern, die anhand des mittlere Abweichungsquadrats der Parzellenrestvarianz geschätzt werden), und die drei Spalten SIG.NIVEAUS, JAHRE mit den Überschriften 88, 89 und 90 geben ihre Richtung und ihre Signifikanzniveaus an. Die Spalte mit D und ND gibt den Unterscheidbarkeitsstatus der beiden Sorten gemäß dem in Teil II Abschnitt 4 beschriebenen Kriterium der 2x1 %-Methode wieder. Die Spalte unter T-BEWERTUNG gibt die veraltete Kenngröße für die T-Bewertung wieder und sollte ignoriert werden.

TGP/8/1 Draft 16: TEIL II: 3: KOMBINIERTES  
 UNTERSCHIEDBARKEITSKRITERIUM ÜBER MEHRERE JAHRE

Seite 77

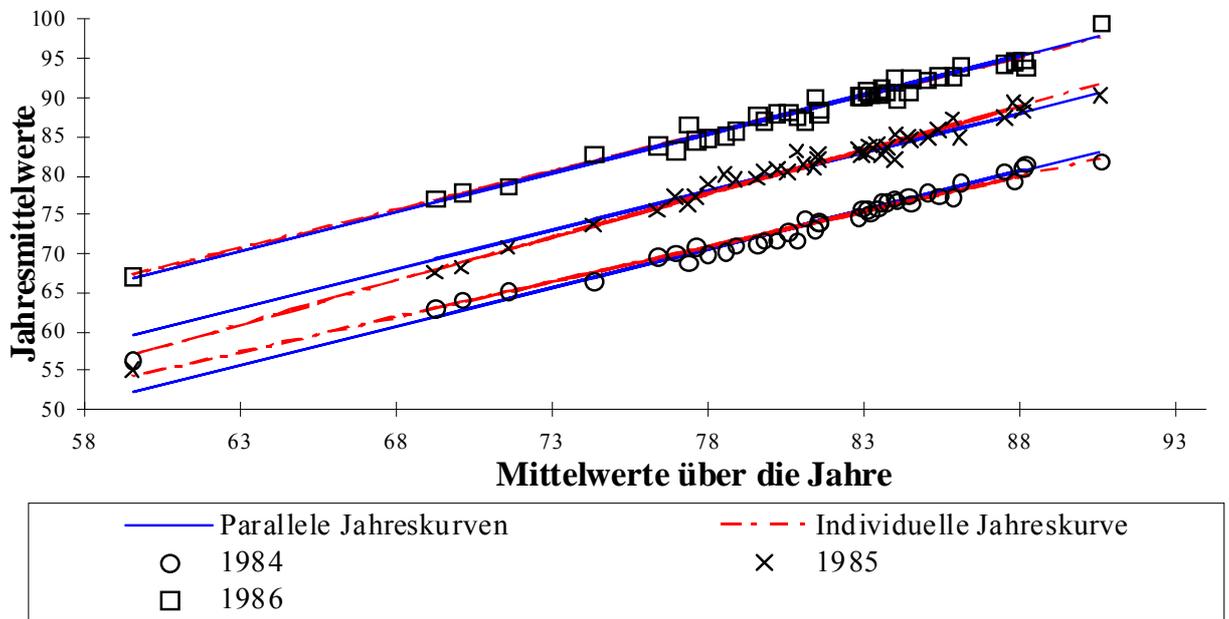
**Tabelle B 3: Ein Beispiel für die Ausgabedaten aus dem COYD-Programm, das den Stand der Unterscheidbarkeit der Kandidatensorten zeigt**

DEUTSCHES WEIDELGRAS (DIPLOID) FRÜH N.I. UPOV 1988-90

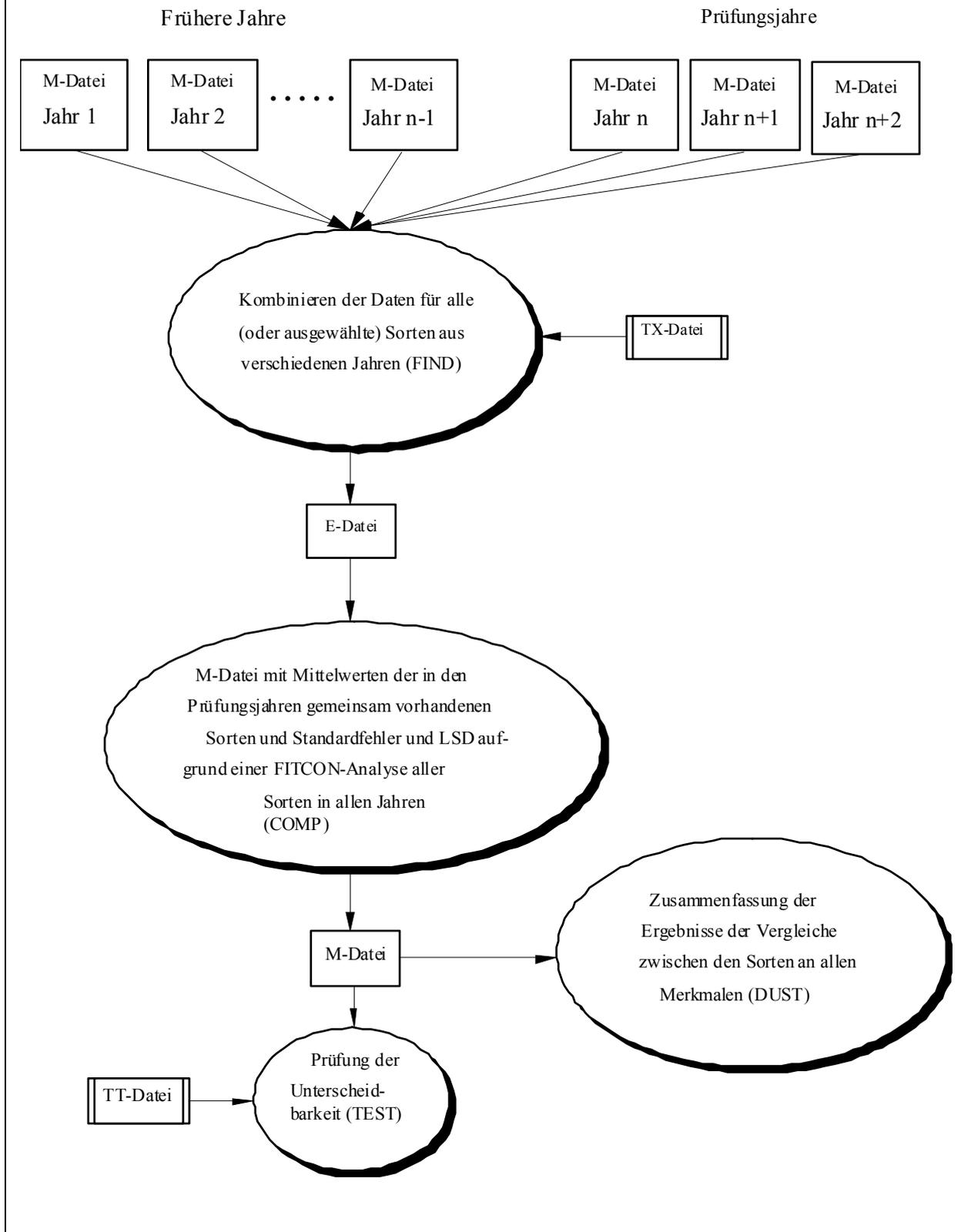
ZUSAMMENFASSUNG FÜR COYD AUF NIVEAU 1,0 %      \*\*\* MIT REGRESS.-BEREINIGUNG, WENN SIGNIFIKANT \*\*\*

KANDIDATENSORTEN		C1	C2	C3	C4	C5	C6	C7	C8	C9
1	R1	D	D	D	D	D	D	D	D	D
2	R2	D	D	D	D	ND	D	D	D	D
3	R3	D	D	D	D	D	D	D	D	D
4	R4	D	D	D	D	D	D	D	D	D
5	R5	D	D	D	D	D	D	D	D	D
6	R6	D	D	D	D	D	D	D	D	D
7	R7	D	D	D	D	D	D	D	D	D
8	R8	D	D	D	D	D	D	D	D	D
9	R9	D	D	D	D	D	D	D	D	D
10	R10	D	D	D	D	D	D	D	D	D
11	R11	D	D	D	D	D	D	D	D	D
12	R1	D	D	D	D	D	D	D	D	D
13	R13	D	D	D	D	ND	D	D	D	D
14	R14	D	D	D	D	D	D	D	D	D
15	R15	D	D	D	D	D	D	D	D	D
16	R16	D	D	D	D	D	D	D	D	D
17	R17	D	D	D	D	D	D	D	D	D
18	R18	D	D	D	D	D	D	D	D	D
19	R19	D	D	D	D	D	D	D	D	D
20	R20	D	D	D	D	D	D	D	D	D
21	R21	D	D	D	D	D	D	D	D	D
22	R22	D	D	D	D	D	D	D	D	D
23	R23	D	D	D	D	D	D	D	D	D
24	R24	D	D	D	D	D	D	D	D	D
25	R25	D	D	D	D	D	D	D	D	D
26	R26	D	D	D	D	D	D	D	D	D
27	R27	D	D	D	D	D	D	D	D	D
28	R28	D	D	D	D	D	D	D	D	D
29	R29	D	D	D	D	D	D	D	D	D
30	R30	D	D	D	D	D	D	D	D	D
31	R31	D	D	D	D	D	D	D	D	D
32	R32	D	D	D	D	D	D	D	D	D
33	R33	D	D	D	D	D	D	D	D	D
34	R34	D	D	D	D	D	D	D	D	D
35	R35	D	D	D	D	D	D	D	D	D
36	R36	D	D	D	ND	D	D	D	D	D
37	R37	D	D	D	D	D	D	D	D	D
38	R38	D	D	D	D	D	D	D	D	D
39	R39	D	D	D	D	D	D	D	D	D
40	R40	D	D	D	D	D	D	D	D	D
41	C1	-	D	D	D	D	D	D	D	D
42	C2	D	-	D	D	D	D	D	D	D
43	C3	D	D	-	D	D	D	ND	D	D
44	C4	D	D	D	-	D	D	D	D	D
45	C5	D	D	D	D	-	D	D	D	D
46	C6	D	D	D	D	D	-	D	D	D
47	C7	D	D	ND	D	D	D	-	D	D
48	C8	D	D	D	D	D	D	D	-	D
49	C9	D	D	D	D	D	D	D	D	-
ANZAHL ND-SORTEN		0	0	1	1	2	0	1	0	0
UNTERSCHIEDBARKEIT		D	D	ND	ND	ND	D	ND	D	D
KANDIDATENSORTE		C1	C2	C3	C4	C5	C6	C7	C8	C9

**Abbildung B1. Jährliche Sortenmittelwerte des Zeitpunkts des Ährenschiebens gegenüber Sortenmittelwerten über die Jahre**



**Abbildung B2. Datenflußdiagramm der Stadien und DUST-Module, die zur Erzeugung der Langzeit-LSD und Durchführung der Langzeit-COYD verwendet werden**



### **3.11 Schemata für die Anwendung von COYD**

3.11.1 Die folgenden vier Fälle stellen im allgemeinen die verschiedenen Situationen dar, die auftreten können, wenn COYD bei der DUS-Prüfung angewandt wird:

Schema A: Die Prüfung wird über zwei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach zwei Wachstumsperioden getroffen (eine Wachstumsperiode könnte ein Jahr sein und wird nachstehend als Periode bezeichnet)

Schema B: Die Prüfung wird über drei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen

Schema C: Die Prüfung wird über drei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen, doch kann eine Sorte auch nach zwei Perioden akzeptiert werden

Schema D: Die Prüfung wird über drei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen, doch kann eine Sorte auch nach zwei Perioden akzeptiert oder zurückgewiesen werden

3.11.2 Die Stadien, in denen die Entscheidungen in den Fällen A bis D getroffen werden, sind in den Abbildungen 1 bis 4 dargestellt. Diese verdeutlichen auch die verschiedenen Standard-Wahrscheinlichkeitsniveaus ( $p_{d2}$ ,  $p_{nd2}$ ,  $p_{d3}$ ,  $p_{u2}$ ,  $p_{nu2}$  und  $p_{u3}$ ), die für die Berechnung des COYD-Kriteriums je nach Fall erforderlich sind. Diese werden wie folgt definiert:

**Wahrscheinlichkeitsniveau verwendet, um zu entscheiden, ob eine Sorte**

$p_{d2}$	nach 2 Perioden unterscheidbar ist
$p_{nd2}$	nach 2 Perioden in einem Merkmal nicht unterscheidbar ist
$p_{d3}$	nach 3 Perioden unterscheidbar ist

3.11.3 In den Abbildungen 1 bis 4 wird das COYD-Kriterium, das beispielsweise anhand des Wahrscheinlichkeitsniveaus  $p_{d2}$  berechnet wird, als  $LSD_{p_{d2}}$  bezeichnet usw. Der Begriff „Unt“ stellt den Unterschied zwischen den Mittelwerten einer Kandidatensorte und einer anderen Sorte für ein Merkmal dar.

3.11.4 Tabelle 1 faßt die verschiedenen Standard-Wahrscheinlichkeitsniveaus zusammen, die für die Berechnung der COYD-Kriterien in jedem der Fälle A bis D erforderlich sind. In Fall B beispielsweise ist nur ein Wahrscheinlichkeitsniveau ( $p_{d3}$ ) erforderlich, während Fall C zwei davon benötigt ( $p_{d2}$ ,  $p_{d3}$ ).

	COYD		
FALL	$p_{d2}$	$p_{nd2}$	$p_{d3}$
A			
B			
C			
D			

Abbildung 1. COYD-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) in Schema A  
COYD

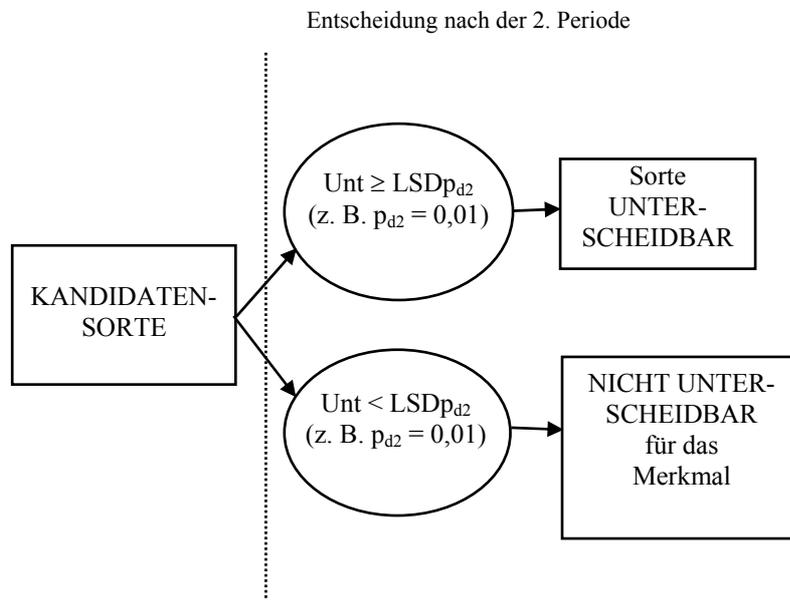
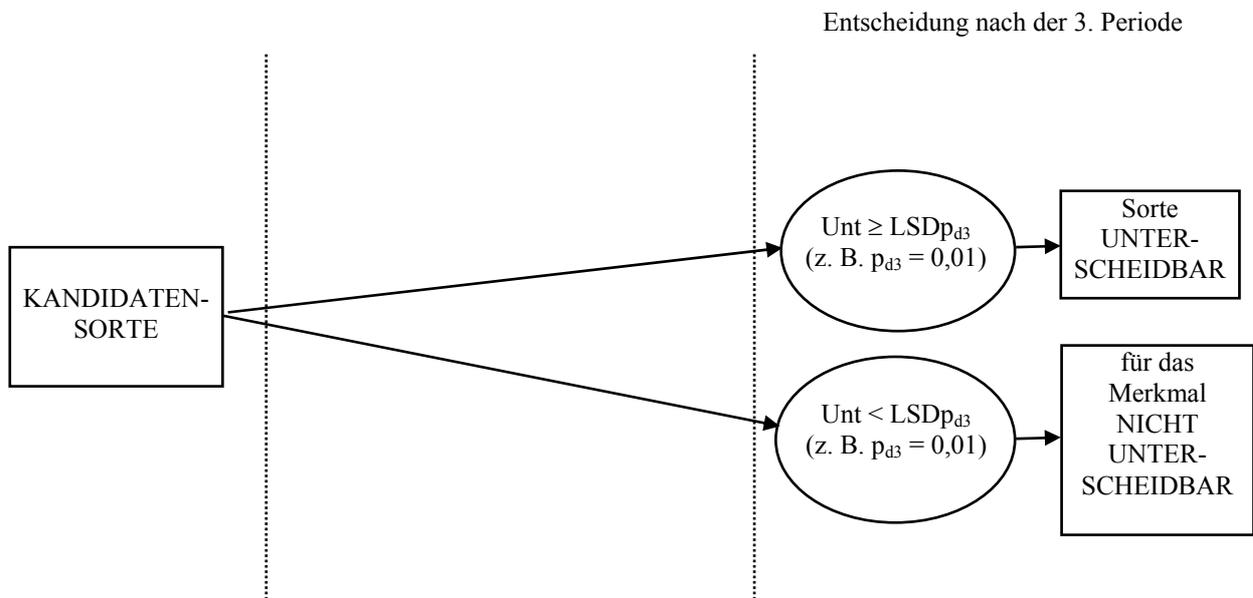


Abbildung 2. COYD-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) in Schema B  
COYD



ANMERKUNG:

„Unt“ ist der Unterschied zwischen den Mittelwerten der Kandidatensorte und einer anderen Sorte für das Merkmal

LSDp ist das COYD-Kriterium, das auf dem Wahrscheinlichkeitsniveau  $p$  berechnet wird.

Abbildung 3. COYD-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) in Schema C  
 COYD

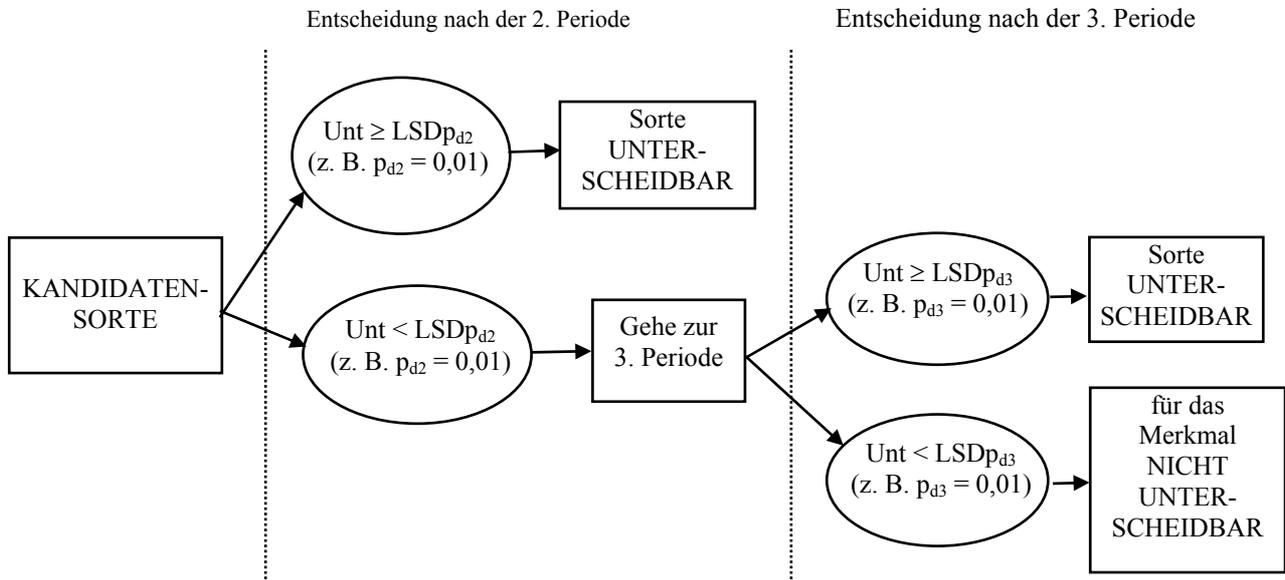
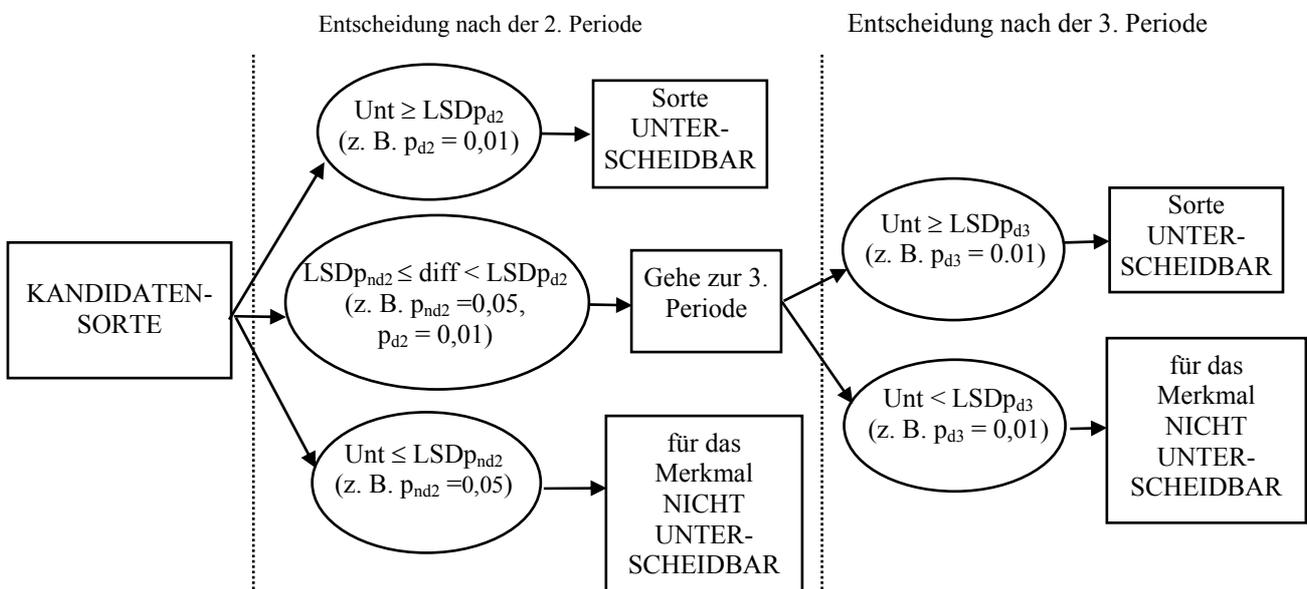


Abbildung 4. COYD-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) in Schema D  
 COYD



ANMERKUNG:

„Unt“ ist der Unterschied zwischen den Mittelwerten der Kandidatensorte und einer anderen Sorte für das Merkmal

LSDp ist das COYD-Kriterium, das auf dem Wahrscheinlichkeitsniveau  $p$  berechnet wird.

## **4. DAS 2X1 %-VERFAHREN**

### **4.1 Voraussetzungen für die Anwendung des Verfahrens**

4.1.1 Das 2x1 %-Kriterium ist ein geeignetes Verfahren für die Prüfung der Unterscheidbarkeit von Sorten, wenn

- das Merkmal quantitativ ist;
- es Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte gibt;
- die Beobachtungen auf Pflanzenbasis (oder Parzellenbasis) über zwei oder mehrere Jahre erfolgen;
- wiederholte Parzellen vorhanden sind.

### **4.2 Das 2x1 %-Kriterium (-verfahren)**

4.2.1 Damit zwei Sorten anhand des 2x1 %-Kriteriums unterscheidbar sind, müssen sie zumindest in zwei von drei Jahren in einem oder mehreren gemessenen Merkmalen in gleicher Richtung auf dem 1 %-Niveau signifikant unterscheidbar sein. Die Tests in jedem Jahr beruhen auf einem zweiseitigen Student-t-Test der Unterschiede zwischen den Sortenmittelwerten mit geschätzten Standardfehlern unter Verwendung des mittleren Abweichungsquadrats des Restes der Analyse der Parzellenmittelwerte Sorte-x-Wiederholung.

4.2.2. Hinsichtlich des 2x1%-Kriteriums im Vergleich zu COYD ist es wichtig anzumerken, daß:

- Informationen verlorengehen, weil sich das Kriterium auf die akkumulierten Entscheidungen stützt, die auf den Ergebnissen der in jedem der Prüfungsjahre durchgeführten t-Tests beruhen. Eine nicht ganz signifikante Differenz auf dem 1 %-Niveau trägt somit nicht mehr zur Trennung eines Sortenpaares bei als eine 0-Differenz oder eine Differenz in entgegengesetzter Richtung. So würden z. B. drei gleichgerichtete Differenzen, wovon eine auf dem 1 %-Niveau und die andere auf dem 5 %-Niveau signifikant ist, nicht als unterscheidbar betrachtet.
- Einige Merkmale sind über die Jahre in der Ausprägung der Unterschiede zwischen Sorten stabiler als andere. Das 2x1 %-Kriterium berücksichtigt jedoch außer der Anforderung, daß die Unterschiede gleichgerichtet sind, nicht die Stabilität bei der Größe der Unterschiede von Jahr zu Jahr.

## 5. DER CHI-QUADRAT-TEST NACH PEARSON MIT KONTINGENZ-TABELLEN

5.1 Eine Kontingenztabelle ist eine Tabelle, die Antworten von Subjekten auf einen Faktor als eine Funktion gegenüber einem anderen Faktor abbilden. Bei der DUS-Prüfung wird sie im allgemeinen für kategoriale Daten verwendet, in denen Individuen einer Sorte zu diskreten Ausprägungsstufen eines Merkmals zugeordnet werden können. Je nach den besonderen Gegebenheiten können verschiedene statistische Tests zur Analyse von Daten in Kontingenztabelle verwendet werden. Der Chi-Quadrat-Test nach Pearson mit Kontingenztabelle ist beispielsweise zweckdienlich, wenn:

- Erfassungen an einem Merkmal zwei oder mehreren Kategorien (Klassen) zugeordnet werden und in einer Kontingenztabelle erfasst werden;
- Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte bestehen;
- die einzige Quelle der Variation durch zufällige Stichproben hervorgerufen wird, z.B. sollte es keine Variation aufgrund der Bodenbeschaffenheit geben, usw.
- der erwartete Mindestwert in jeder Kategorie sollte fünf sein.

5.2 In einigen Fällen kann die Unterscheidbarkeit jedoch durch die Klassifikation einzelner Sorten in umfassende Gruppen und durch den Nachweis statistisch verschiedener Gruppenmuster für verschiedene Sorten begründet werden. Derartige Beispiele sind Zählungen aufgrund der Blütenfarbgruppen – wie dunkel blauviolett gegenüber nicht dunkel blauviolett und die Klassen für Krankheits-/Schadorganismen-/Nematodenbefall. Die auf Zählungen von Individuen in einer Stichprobe/Population, die jeder von mehreren Klassen angehören, beruhenden Daten setzen statistischen Analysen voraus, die für kategoriale Daten geeignet sind.

5.3 Für die Anwendung der Chi-Quadrat-Analyse für Züchterrechtszwecke sollte geprüft werden, wie bestimmte Schlußfolgerungen bezüglich der Unterscheidbarkeit erzielt werden, indem bestimmte Hypothesen anhand der Klassifikationsdaten formuliert werden.

Die in dieser Analyse angewandte Standardformel für die Chi-Quadrat-Kenngröße ist:

$$X^2 = \sum \frac{(\text{erfaßter Wert einer Klasse} - \text{erwarteter Wert einer Klasse})^2}{\text{erwarteter Wert}}$$

5.4 Die Chi-Quadrat-Verteilung ist demzufolge eine kontinuierliche Verteilung aufgrund einer zugrundeliegenden Normalverteilung.

5.5 Vor der Anwendung des Chi-Quadrat-Tests sind die nachstehenden Vorsichtsmaßnahmen zu bedenken:

- 1) Die Wahl der zu prüfenden Hypothese sollte auf vorher bekannten Fakten oder Grundsätzen basieren.
- 2) Wenn die Hypothese gegeben ist, solle es möglich sein, die erwarteten Werte für jede Klasse richtig zuzuweisen. Die Anwendung des Chi-Quadrat-Tests ist zu

vermeiden, wenn die geringste erwartete Klasse weniger als fünf beträgt. Durch Erhöhung der Stichprobengröße kann die Größe der kleinsten erwarteten Klasse erhöht werden. Wenn einige Klassen eine Größe von weniger als fünf haben, sind alternativ entweder diese angrenzenden Klassen zusammenzulegen, um die Größe der zusammengelegten Klasse auf fünf oder mehr als fünf zu erhöhen, oder es ist ein exakter Test anzuwenden.

3) Die Freiheitsgrade werden als die Anzahl Klassen definiert, die unabhängig sind, damit ihnen ein willkürlicher Wert zugeteilt werden kann. Wenn beispielsweise zwei Klassen vorhanden sind, betragen die Freiheitsgrade  $2-1 = 1$ . Demzufolge beträgt die Zahl der Freiheitsgrade für den Chi-Quadrat-Test bei der Anwendung dieses Verfahren für die Prüfung einer Hypothese eins weniger als die Anzahl Klassen.

4) Situationen mit zwei Klassen, die eher einer Binomialverteilung folgen, wobei  $np$  oder  $nq$  unter 5 liegen, sind zu vermeiden. In solchen Situationen sind die erwarteten Werte anhand von Formeln zu berechnen, die auf der Binomialverteilung beruhen. In einer Situation mit zwei Klassen entspricht  $np$  der Größe einer der Klassen, die durch die Anzahl Ereignisse ( $n$ )  $\times$  die Wahrscheinlichkeit, in eine dieser Klassen zu fallen ( $p$ ), bestimmt wird. Gleichermaßen wird die Größe der anderen Klasse ( $nq$ ) durch  $n$ -Mal die Wahrscheinlichkeit ( $q$ ) bestimmt, in diese Klasse zu fallen. Somit ist in einer Situation, in der die Wahrscheinlichkeit, in eine der beiden Klassen zu fallen, gleich ( $p=q=0,5$ ) und die Stichprobengröße 10 ( $n$ ) ist, die erwartete Anzahl in jeder Klasse 5. Für die Bestimmung des Chi-Quadrat-Tests mit nur einem Freiheitsgrad ist stets die Yates-Bereinigung anzuwenden.

5.6 Prüfen wir folgende Daten für die Benotung der Krankheit zweier Generationen einer Kandidatensorte von Luzerne und vier allgemein bekannte Sorten. Die benotete Krankheit war *Colletotrichum trifolii* (Merkmal 19, TG/6/5, Luzerne). Die Benotung erfolgte auf der Klassenskala 5, wobei Klasse 1 (Note 9) resistent und Klasse 5 (Note 1) anfällig ist.

*Anzahl Pflanzen, die in verschiedenen Klassen bei jeder Sorte nach 7-10 Tagen Inokulation gezählt werden*

Note (Klasse)	Kandidat		Sorte 1	Sorte 2	Sorte 3	Sorte 4
9(1)	34		12	6	1	7
7(2)	4		7	6	5	10
5(3)	1		9	5	5	5
3(4)	1		7	9	8	7
1(5)	6		9	19	9	15
Insgesamt	46		44	45	28	44

5.7 Aus der Tabelle ist zu ersehen, daß die beiden Generationen der Kandidatensorte mehr Pflanzen in der resistenten Kategorie als die vier allgemein bekannten Sorten haben. Um die Signifikanz dieser Sortenunterschiede statistisch zu prüfen, sind jedoch zwei Hypothesen zu formulieren:

1) Ob die vier allgemein bekannten Sorten von der Kandidatensorte in der Verteilung der Punktzahlen, d. h. anhand der Prüfung der Nullhypothese, signifikant verschieden sind. Die Nullhypothese lautet in diesem Fall, daß alle Sorten eine ähnliche Reaktion auf die Kronwurzel von *Colletotrichum* zeigen. Dies kann nur anhand der Prüfung der „ $\chi^2$ -Unterscheidbarkeit“ erfolgen.

5.8 Um den für die Anwendung des Chi-Quadrat-Tests erforderlichen Mindestwert zu erreichen, ist es notwendig, Klassen in Form von intermediär zusammengelegten Klassen zusammenzulegen.

Die erfaßten Daten sind nun reduziert auf:

Klasse/Punktzahl	Kandidat	Sorte 1	Sorte 2	Sorte 3	Sorte 4
1	34	12	6	1	7
2	6	23	20	18	22
3	6	9	19	9	15
Insgesamt	46	44	45	28	44

5.9 Für jeden Vergleich einer Kandidatensorte mit jeder allgemein bekannten Sorte **Error! Bookmark not defined.** wird eine zweiachsige Tabelle der erfassten Werte gebildet. Die erwarteten Werte werden als Produkt der Gesamtwerte von Reihe und Spalte geteilt durch den Gesamtwert berechnet und die Chi-Quadrat Statistik wird berechnet. Die Verteilung der erwarteten Werte für verschiedene Sorten ist wie folgt:

Erfasst für Sorte 1

Klasse/Punktzahl	Kandidat	Sorte 1	Insgesamt
1	34	12	46
2	6	23	29
3	6	9	15
<b>Insgesamt</b>	<b>46</b>	<b>44</b>	<b>90</b>

Erwartet für Sorte 1

Klasse/Punktzahl	Kandidat	Sorte 1	Insgesamt
1	$23.5=46 \times 46 / 90$	$22.5=46 \times 44 / 90$	46
2	$14.8=29 \times 46 / 90$	$14.2=29 \times 44 / 90$	29
3	$7.7=15 \times 46 / 90$	$7.3=15 \times 44 / 90$	15
Insgesamt	46	44	90

Gleichermaßen sind die erwarteten Werte bei Anwendung der Tabelle der erfassten Daten in 5.3.8 für die **Error! Bookmark not defined.**Sorten 2,3 und 4:

Klasse/Punktzahl	Kandidat	<b>Error! Bookmark not defined.Sorte 2</b>	Insgesamt
1	20,2	19,8	40
2	13,1	12,9	26

3	12,6	12,4	25
Insgesamt	46	45	91

Klasse/Punktzahl	Kandidat	Error! Bookmark not defined.Sorte 3	Insgesamt
1	21,8	13,2	35
2	14,9	9,1	24
3	9,3	5,7	15
Insgesamt	46	28	74

Klasse/Punktzahl	Kandidat	Error! Bookmark not defined.Sorte 4	Insgesamt
1	21,0	20,0	41
2	14,3	13,7	28
3	10,7	10,3	21
Insgesamt	46	44	90

5.10 Für die Berechnung der „ $\chi^2$ -Unterscheidbarkeit“ für die **Error! Bookmark not defined.Sorte 1** ist

$$\chi^2 = (34-23.5)^2/23.5 + (12-22.5)^2/22.5 + (6-14.8)^2/14.8 + (23-14.2)^2/14.2 + (6-7.7)^2/7.7 + (9-7.3)^2/7.3 = 21.1$$

mit (Keine Reihe – 1)(Kein Punkt – 1) = 2 df

5.11 Die Anzahl Freiheitsgrade für die Suche auf der  $\chi^2$ -Tabelle beträgt eins weniger als die Anzahl Reihen multipliziert mit der Anzahl Spalten, d. h. (3 – 1) x (2-1) = 2.

5.12 Bei P = 0,01 für 2 df beträgt der Tabellenwert 9,21. Die berechnete  $\chi^2$ -Unterscheidbarkeit ist geringer als der  $\chi^2$ -Tabellenwert. Demzufolge wird die Nullhypothese zurückgewiesen, daß die **Error! Bookmark not defined.Sorte 1** eine ähnliche Reaktion auf die Krankheit zeigt wie diejenige der Kandidatensorte.

5.13 Gleichmaßen ist die berechnete „ $\chi^2$ -Unterscheidbarkeit“ für die **Error! Bookmark not defined.Sorte 2**, die **Error! Bookmark not defined.Sorte 3** und die **Error! Bookmark not defined.Sorte 4** 33,9, 35,4 bzw. 30,8; alle sind höher als der  $\chi^2$ -Tabellenwert von 9,21 bei 2 df.

5.14 Demzufolge sind alle vier allgemein bekannten Sorten von der Kandidatensorte in bezug auf die Reaktion auf die Kronwurzeln von Colletotrichum signifikant verschieden.

## 6. EXAKTER FISHER-TEST

Der exakte Fisher-Test ist ein statistischer Test, der bei der Analyse kategorischer (qualitativer) Daten angewandt wird, wenn die Anzahl Stichproben (d. h. die Stichprobengröße) gering ist. Der exakte Fisher-Test mit 2x2-Kontingenztabelle ist zweckdienlich, wenn:

- Erfassungen an einem Merkmal zwei oder mehreren Kategorien (Klassen) zugeordnet werden;
- die einzige Quelle der Variation durch zufällige Stichproben hervorgerufen wird, z.B. sollte es keine Variation aufgrund der Bodenbeschaffenheit geben, usw.
- die erwarteten Werte in jeder Kategorie weniger als 10 betragen.

### 6.1 Prüfung der Unterscheidbarkeit

6.1.1 Der exakte Fisher-Test wird angewandt, um zu bestimmen, ob es nicht zufallsbedingte Assoziationen zwischen zwei kategorischen Variablen in einer 2x2-Kontingenztabelle<sup>5</sup> gibt, und kann angewandt werden, wenn die Anzahl Stichproben für eine oder mehrere Kategorien für jede Sorte weniger als 10 beträgt (vergleiche die Zellen mit fettgedruckten Rahmen in Tabelle 1) oder wenn die Tabelle stark asymmetrisch ist. Wenn eine große Anzahl Stichproben vorliegt (d. h. 10 oder mehr), wird häufig ein Chi-Quadrat-Test vorgezogen.

6.1.2 Dieser Test gilt lediglich für die Analyse kategorischer Daten. Die nachstehenden hypothetischen Beispiele verdeutlichen dieses Verfahren:

#### Beispiel 1

6.1.3 Im folgenden Beispiel wird die Häufigkeit dunkelblauer Blüten als maßgebendes Merkmal bei der DUS-Prüfung verwendet. In diesem Beispiel eines DUS-Anbauversuchs mit zwei Sorten werden die Pflanzen so benotet, daß sie dunkelblaue Blüten haben oder keine dunkelblauen Blüten haben.

6.1.4 Nehmen wir an, daß die beiden Sorten (Sorte 1 und Sorte 2) im Anteil der dunkelblauen Blüten einige erfaßte Unterschiede aufweisen. Die Prüfer müssen in der Lage sein, zuverlässig zu bestimmen, ob diese Unterschiede als deutlich erkennbar akzeptiert werden können. Der exakte Fisher-Test bietet ein akzeptiertes Verfahren für die Prüfung der Hypothese, daß die erfaßten Unterschiede statistisch signifikant sind. In Tabelle 1 sind hypothetische Daten aus insgesamt 24 Pflanzen angegeben.

Tabelle 1: Eine 2x2-Kontingenztabelle – Anzahl Pflanzen mit nicht dunkelblauen und mit dunkelblauen Blüten, die bei Sorte 1 und Sorte 2 erfaßt wurden

+	Sorte 1	Sorte 2	Insgesamt
Nicht dunkelblau	<b>4</b>	<b>9</b>	13
Dunkelblau	<b>8</b>	<b>3</b>	11
Insgesamt	12	12	24

<sup>5</sup> Eine Kontingenztabelle wird benutzt, um die Beziehung zwischen zwei oder mehreren Variablen, zumeist kategorischen Daten, zu erfassen und zu analysieren.

In einer 2x2-Kontingenztabelle ist die Zahl der Freiheitsgrade stets 1.

6.1.5 Welches ist die Wahrscheinlichkeit, daß Sorte 1 von Sorte 2 aufgrund dieses Merkmals unterscheidbar ist, wenn bekannt ist, daß 11 dieser 24 Blüten dunkelblau sind und 8 davon von Sorte 1 und 3 von Sorte 2 stammen? Oder mit anderen Worten, ist der erfaßte Unterschied bei der Blütenfarbe mit den Sortenunterschieden verbunden, oder ist es wahrscheinlich, daß er sich aus einer zufälligen Stichprobenerhebung ergeben hat? Das Fisher-Verfahren berechnet die genaue Wahrscheinlichkeit einer nicht zufälligen Assoziation aus einer 2x2-Kontingenztabelle anhand einer hypergeometrischen Verteilung<sup>6</sup>. In diesem Fall wird die Wahrscheinlichkeit als die Summe der Wahrscheinlichkeiten für jedes mögliche Auftreten berechnet, das größer als oder gleich groß wie das erfasste ist. Folglich würde die Anzahl erfolgreich ausgebildeter dunkelblauer Blüten zusätzlich zu den beobachteten 9, 10 oder 11 für Sorte 1 sowie 2, 1 oder 0 für Sorte 2 betragen.

6.1.6 Wenn die obigen Zellen mit einer algebraischen Notation versehen werden, wird die allgemeine Formel für die Berechnung der Wahrscheinlichkeit der erfaßten Zahlen ermittelt (Tabelle 2).

*Tabelle 2: Algebraische Notation für den exakten Fisher-Test*

	Sorte 1	Sorte 2	Insgesamt
Nicht dunkelblau	a	b	a + b
Dunkelblau	c	d	c + d
Insgesamt	a + c	b + d	n

$$p = \frac{(a+b)! (c+d)! (a+c)!(b+d)!}{n!a!b!c!d!}$$

6.1.7 wobei p die Wahrscheinlichkeit nach dem exakten Fisher-Test ist, eine nicht zufällige Verteilung zwischen den Sorten und den Merkmalen zu ermitteln (! ist das Symbol für Fakultät).

6.1.8 Wenn die algebraischen Notationen in Tabelle 2 durch die erfaßte Anzahl aus Tabelle 1 ersetzt werden:

$$p = \frac{(13)! (11)! (12)!(12)!}{24!4!9!8!3!}$$

Nach der Auflösung der Fakultäten:

$$p = 0,05$$

6.1.9 Die Interpretation des vom exakten Fisher-Test berechneten p-Wertes ist einfach. Im obigen Beispiel  $p = 0,05$ , was bedeutet, daß die Wahrscheinlichkeit, daß bei der Stichprobengröße und der Verteilung in Tabelle 1 die erfaßten Unterschiede allein auf die Stichprobenerhebung zurückzuführen sind, 5 % beträgt. Angesichts der geringen Stichprobengröße und der Notwendigkeit, daß die Sorten deutlich erkennbar sind, steht es den Prüfungsbehörden frei,  $p = 0,01$  als die Obergrenze für die Signifikanzakzeptierbarkeit

<sup>6</sup> Eine hypergeometrische Verteilung ist eine diskrete Wahrscheinlichkeitsverteilung, die die Anzahl Erfolge in einer Sequenz von Ziehungen aus einer begrenzten Population ohne Ersetzung beschreibt.

unserer Nullhypothese zu wählen. Somit würde eine Prüfungsbehörde aus diesem Beispiel folgern, daß der beim dunkelblauen kontra nicht dunkelblauen Merkmal erfaßte Unterschied nicht signifikant verschieden ist und die beiden Sorten (Sorte 1 und Sorte 2) auf dieser Grundlage nicht unterscheidbar sind.

Beispiel 2

6.1.10 Die Erfassungen für die Sorte 3 und die Sorte 4 für dasselbe Merkmal und die Erfassungen sind in Tabelle 3 wiedergegeben:

Tabelle 3: Anzahl Pflanzen mit nicht dunkelblauen und dunkelblauen Blüten, die bei Sorte 3 und Sorte 4 erfaßt wurden

	Sorte 3	Sorte 4	Insgesamt
Nicht dunkelblau	1	9	10
Dunkelblau	11	3	14
Insgesamt	12	12	24

Wenn die obigen Werte in die hypergeometrische Fisher-Verteilung gesetzt werden:

$$p = \frac{(10!) (14!) (12!) (12)!}{24! 1! 9! 11! 3!}$$

Nach der Auflösung der Fakultäten wird der Fisher-Wahrscheinlichkeitswert berechnet als:

$$p = 0,001$$

6.1.11 In diesem besonderen Fall wird die Nullhypothese (daß die Sorten aufgrund des Merkmals dunkelblau kontra nicht dunkelblau ähnlich sind) zurückgewiesen, weil die berechnete Fisher-Wahrscheinlichkeit weit geringer als das zulässige Signifikanzniveau ( $p = 0,01$ ) ist. Dementsprechend sollten die beiden Sorten (Sorte 3 und Sorte 4) für unterscheidbar erklärt werden.

## 7. MATCH-ANSATZ

### 7.1 Voraussetzungen für die Anwendung der Methode

7.1.1 Die Match-Methode ist ein geeignetes Verfahren für die Prüfung der Unterscheidbarkeit von Sorten, wenn:

- Daten aus mehr als einem Jahr analysiert werden,
- Erfassungen auf Pflanzenbasis (oder Parzellenbasis) im zweiten Jahr mit Erfassungen vom Züchter im ersten Jahr verglichen werden,
- Unterschiede zwischen Pflanzen (oder Parzellen) einer Sorte aufgrund von Informationen aus einem Versuch im ersten Jahr angegeben werden,
- die Voraussetzungen der Methode hängen vom besonderen statistischen Test ab, der verwendet wird, (z.B. LSD, Mehrfachreihentest (MRT), Chi-Quadrat oder exakter Fisher-Test).

### 7.2 Match-Methode

7.2.1 Die Match-Methode zur Prüfung der Unterscheidbarkeit wurde für die Analyse von Daten aus mehr als einer Wachstumsperiode entwickelt. Die Versuche werden in der ersten Wachstumsperiode vom Züchter und in der zweiten Wachstumsperiode von der Prüfungsbehörde durchgeführt (vergleiche Dokument TGP/6 Abschnitt 2.1). Ob die Unterschiede hinreichend stabil sind, wird anhand eines statistischen Tests geprüft (z.B. LSD, Mehrfachreihentest (MRT), Chi-Quadrat oder exakter Fisher-Test) um zu beurteilen, ob die Unterschiede bei den Sortenmittelwerten im zweiten Jahr signifikant sind und mit der „Richtung der Unterschiede“ übereinstimmen, die von den Züchtern im ersten Jahr gemeldet werden. Die Wahl der statistischen Verfahren hängt vom Ausprägungstyp des entsprechenden Merkmals ab. Damit zwei Sorten nach der Match-Methode unterscheidbar sind, müssen sich die Sorten signifikant in dieselbe Richtung wie vom Züchter gemeldet unterscheiden.

7.2.2 Die Voraussetzungen der Methode hängen vom besonderen statistischen Verfahren ab, das angewendet wird (z.B. LSD, Mehrfachreihentest (MRT), Chi-Quadrat oder exakter Fisher-Test). Für quantitative Merkmale kann der Test beispielsweise auf einer zweiseitigen LDS beruhen, wenn es eine Kandidatensorte, oder auf einem Mehrfachreihentest (MRT), wenn mehr als eine Kandidatensorte im Anbauversuch enthalten ist. Ein Chi-Quadrat-Test oder ein exakter Fisher-Test können für pseudo-qualitative oder qualitative Merkmale verwendet werden, wenn die Voraussetzungen für diese Tests gegeben sind. Obwohl diese Tests bei Anbauversuchen für fremdbefruchtende Sorten am nützlichsten sind, können sie jedoch gleichermaßen auf Anbauversuche mit selbstbefruchtenden und vegetativ vermehrten Sorten angewandt werden, sofern die entsprechenden Kriterien erfüllt sind.

7.2.3 Die Match-Methode betrifft in der Regel relativ begrenzte Versuche bei denen die Anzahl von Sorten **Error! Bookmark not defined.** im Versuch auf die Kandidatensorten und die ähnlichsten allgemein bekannten Sorten beschränkt ist.

## **8. VERFAHREN FÜR DIE PRÜFUNG DER HOMOGENITÄT ANHAND VON ABWEICHERN**

### **8.1 Fester Populationsstandard**

#### **8.1.1 Einleitung**

Dokument TGP/10 Abschnitt 4 gibt Anleitung, wann es geeignet ist, das Verfahren der Prüfung der Homogenität anhand von Abweichern unter Verwendung eines festen Populationsstandards anzuwenden. Es gibt ferner Anleitung bezüglich der Bestimmung der pflanzenabhängigen Einzelheiten wie der Stichprobengröße und der zulässigen Anzahl Abweicher. Dieser Abschnitt beschreibt das Abweicherverfahren aus folgenden Perspektiven:

- Anwendung des Abweicherverfahrens zur Prüfung der Homogenität bei einer Pflanze.
- die zu prüfenden Aspekte bei der Entscheidung über die pflanzenabhängigen Einzelheiten für die Prüfung der Homogenität einer Pflanze anhand des Abweicherverfahrens. Zu diesen Einzelheiten gehören die Stichprobengröße, die zulässige Anzahl Abweicher, ob in mehr als einem Jahr geprüft werden soll und ob die Sequenzprüfung durchgeführt werden soll.

#### **8.1.2 Anwendung des Verfahrens zur Prüfung der Homogenität bei einer Pflanzenart**

8.1.2.1 Um das Verfahren für die Prüfung der Homogenität bei einer Pflanzenart anzuwenden, werden folgende artabhängigen Einzelheiten entweder den UPOV-Prüfungsrichtlinien entnommen oder aufgrund der Erfahrung entschieden, insbesondere mit Bezug auf andere UPOV-Prüfungsrichtlinien für vergleichbare Sortentypen:

- eine Stichprobengröße, z. B. 100 Pflanzen
- eine Höchstzahl von Abweichern, die in der Stichprobe zulässig ist, z. B. 3
- ein fester Populationsstandard, z. B. 1 %
- und eine Akzeptanzwahrscheinlichkeit, z. B. mindestens 95 %

8.1.2.2 Als nächstes wird eine Stichprobe der Kandidatensorte basierend auf der richtigen Pflanzenanzahl betrachtet und die Anzahl Abweicher gezählt. Wenn die Anzahl geringer als oder gleich der zulässigen Höchstzahl ist, wird die Sorte als homogen akzeptiert, ansonsten wird sie als nichthomogen zurückgewiesen. Bei diesen Entscheidungen könnten zwei statistische Fehler auftreten. Die Risiken, diese Fehler zu begehen, werden durch die Wahl der Stichprobengröße und der zulässigen Höchstzahl von Abweichern unter Kontrolle gehalten.

8.1.2.3 Der feste Populationsstandard oder „Populationsstandard“ ist der maximale Prozentsatz der Abweicher, der zulässig wäre, wenn alle Individuen der Sorte geprüft werden könnten. Im obigen Beispiel beträgt er 1 %. Sorten mit weniger als dem Populationsstandard der Abweicher sind homogen, und diejenigen mit mehr als dem Populationsstandard sind nicht homogen. Es können jedoch nicht alle Individuen der Sorte geprüft werden; statt dessen muß eine Stichprobe geprüft werden.

8.1.2.4 Nehmen wir eine Sorte, die, wenn alle Individuen der Sorte geprüft würden, nicht mehr als den Populationsstandard der Abweicher hätte. Bei der Stichprobenerhebung gibt es zwei mögliche Ergebnisse. Entweder enthält die Stichprobe nicht mehr als die zulässige Höchstzahl an Abweichern; in diesem Fall wird die Sorte als homogen akzeptiert, oder die Stichprobe enthält mehr als die zulässige Höchstzahl Abweicher, und die Sorte wird

zurückgewiesen. Im letzteren Fall wäre ein statistischer Fehler mit der Bezeichnung „Fehler vom Typ I“ aufgetreten. Die Wahrscheinlichkeit, diese Sorte zu akzeptieren, und die Wahrscheinlichkeit eines Fehlers vom Typ I sind wie folgt verknüpft:

$$\text{„Akzeptanzwahrscheinlichkeit“} + \text{„Wahrscheinlichkeit eines Fehlers vom Typ I“} = 100 \%$$

8.1.2.5 Die Wahrscheinlichkeit, eine Sorte aufgrund einer Stichprobe zu akzeptieren oder zurückzuweisen, hängt von der Stichprobengröße, der zulässigen Höchstzahl Abweicher und dem Prozentsatz der Abweicher ab, die festgestellt würden, wenn alle Individuen der Sorte geprüft würden. Die Stichprobengröße und die zulässige Höchstzahl Abweicher werden so gewählt, daß sie die „Akzeptanzwahrscheinlichkeit“ erfüllen, die die Mindestwahrscheinlichkeit ist, daß eine Sorte mit dem Populationsstandard von Abweichern akzeptiert wird. So wurden im Beispiel oben die Stichprobengröße und die Höchstzahl Abweicher so gewählt, daß sie eine mindestens eine Wahrscheinlichkeit von 95 % ergeben, daß eine Sorte akzeptiert wird, die, wenn alle Individuen der Sorte geprüft würden, 1 % Abweicher hätte.

8.1.2.6 Zur Überprüfung der Stichprobengröße und der Höchstzahl Abweicher im obigen Beispiel sollte die Tabelle A geprüft werden, die die Tabelle 5 und die Abbildung 5 als relevant für einen Populationsstandard von 1 % und eine Akzeptanzwahrscheinlichkeit von  $\geq 95 \%$  aufführt. Bei der Betrachtung der Tabelle 5 ist festzustellen, daß eine Stichprobengröße von 100 (zwischen 83 und 137) und eine Höchstzahl von Abweichern von 3 eine Akzeptanzwahrscheinlichkeit von  $> 95 \%$  für einen Populationsstandard von 1 % ergibt. Die Abbildung 5 ist detaillierter: Die unterste der vier Kurven gibt die Wahrscheinlichkeit eines Fehlers vom Typ I für die verschiedenen Stichprobengrößen und Höchstzahlen der Abweicher an, die in Tabelle 5 aufgeführt sind. Für einen Populationsstandard von 1 %, eine Stichprobengröße von 100 und die Zulässigkeit von bis zu 3 Abweichern beträgt die Wahrscheinlichkeit eines Fehlers vom Typ I daher 2 %, so daß die Wahrscheinlichkeit, daß aufgrund einer solchen Stichprobe eine Sorte mit dem Populationsstandard von Abweichern, d. h. 1 %,  $100 \% - 2 \% = 98 \%$ , beträgt, akzeptiert wird, was höher als die erforderliche „Akzeptanzwahrscheinlichkeit“ (95 %) ist.

8.1.2.7 Aus Abbildung 5 ist zu ersehen, daß die Wahrscheinlichkeit eines Fehlers vom Typ I in dem Maße zunimmt, wie die Stichprobengröße zunimmt, und die Wahrscheinlichkeit, eine Sorte mit dem Populationsstandard, d. h. 1 %, von Abweichern abnimmt, bis diese Wahrscheinlichkeit zu gering wird, um die „Akzeptanzwahrscheinlichkeit“ zu erfüllen, und es notwendig wird, die Höchstzahl der Abweicher gemäß Tabelle 5 zu erhöhen.

8.1.2.8 Ebenso wie eine Sorte mit dem Populationsstandard oder weniger Abweichern aufgrund einer Stichprobe entweder akzeptiert oder zurückgewiesen (Fehler vom Typ I) werden kann, kann eine Sorte mit mehr als dem Populationsstandard der Abweicher entweder akzeptiert oder zurückgewiesen werden. Eine Sorte mit mehr als dem Populationsstandard von Abweichern aufgrund einer Stichprobe zu akzeptieren, wird als „Fehler vom Typ II“ bezeichnet. Die Wahrscheinlichkeit eines Fehlers vom Typ II hängt davon ab, wie nichthomogen die Sorte ist. Die drei oberen Kurven in Abbildung 5 geben die Wahrscheinlichkeiten von Fehlern vom Typ II für drei Nichthomogenitätsgrade für die verschiedenen Stichprobengrößen und Höchstzahlen von Abweichern an, die in Tabelle 5 aufgeführt sind. Die drei Nichthomogenitätsgrade sind 2, 5 und 10 Mal der Populationsstandard. Sie werden durch die oberste, mittlere und unterste der drei oberen Kurven dargestellt. Somit beträgt bei einer Stichprobengröße von 100 und bis zu 3 zulässigen Abweichern die Wahrscheinlichkeit, eine Sorte mit 2 % Abweichern zu akzeptieren, 86 %,

eine Sorte mit 5 % Abweichern zu akzeptieren, 26 % und eine Sorte mit 10 % Abweichern zu akzeptieren, 1 %. Im allgemeinen

- ist die Wahrscheinlichkeit eines Fehlers vom Typ II um so geringer, je größer die Nichthomogenität ist.
- nimmt für eine gegebene Höchstzahl Abweicher die Wahrscheinlichkeit eines Fehlers vom Typ II in dem Maße ab, wie die Stichprobengröße zunimmt.
- nimmt die Wahrscheinlichkeit eines Fehlers vom Typ II in dem Maße zu, wie die Höchstzahl der Abweicher zunimmt.

8.1.3 Bei der Entscheidung über die Anwendung des Verfahrens zu berücksichtigende Aspekte

8.1.3.1 Im vorhergehenden Abschnitt wurde festgestellt, daß die Wahrscheinlichkeit, eine Sorte mit dem Populationsstandard oder weniger Abweichern zu akzeptieren oder zurückzuweisen (Fehler vom Typ I), und die Wahrscheinlichkeit, eine Sorte mit mehr als dem Populationsstandard von Abweichern zu akzeptieren (Fehler vom Typ II) oder zurückzuweisen, von der Wahl der Stichprobengröße und der zulässigen Höchstzahl von Abweichern abhängt. Der Rest dieses Kapitels behandelt die Frage, wie diese Wahl genutzt werden kann, um die Risiken der Fehler vom Typ I und vom Typ II auszugleichen. Dies wird an einer Reihe von Beispielen verdeutlicht. Die Erörterung wird erweitert, um die Situation einzubeziehen, in der die Prüfung über mehr als ein Jahr durchgeführt wird, einschließlich der Möglichkeit, Sequenzprüfungen durchzuführen, um den Aufwand der Stichprobenerhebung zu minimieren. Es werden Tabellen und Abbildungen vorgelegt, denen die Wahrscheinlichkeiten von Fehlern vom Typ I und vom Typ II oder verschiedene Kombinationen von Populationsstandards und Akzeptanzwahrscheinlichkeiten zu entnehmen sind. Zudem werden Einzelheiten über die Art und Weise der direkten Berechnung der Wahrscheinlichkeiten sowohl für Prüfungen in einem einzigen Jahre als auch in zwei oder mehreren Jahren, einschließlich der Zweiphasen-Prüfung, angegeben.

8.1.3.2 Die beiden oben beschriebenen Fehlertypen lassen sich in die nachstehende Tabelle zusammenfassen:

<b>Entscheidung, die getroffen würde, wenn alle Pflanzen einer Sorte geprüft werden könnten</b>	Entscheidung aufgrund der Anzahl Abweicher in einer Stichprobe	
	die Sorte wird als homogen akzeptiert	die Sorte wird als nicht homogen zurückgewiesen
die Sorte ist homogen	gleiche Entscheidung	andere Entscheidung, Fehler vom Typ I
die Sorte ist nicht homogen	andere Entscheidung, Fehler vom Typ II	gleiche Entscheidung

8.1.3.3 Die Wahrscheinlichkeit von Fehlern vom Typ II hängt davon ab, „wie nichthomogen“ die Kandidatensorte ist. Wenn sie viel homogener ist als der Populationsstandard ist, dann ist die Wahrscheinlichkeit eines Fehlers vom Typ II gering und die Wahrscheinlichkeit, eine solche Sorte zu akzeptieren, gering. Wenn andererseits die

Kandidatensorte nur geringfügig nichthomogener als der Standard ist, dann ist die Wahrscheinlichkeit eines Fehlers vom Typ II hoch. Die Wahrscheinlichkeit der Akzeptanz nähert sich der Akzeptanzwahrscheinlichkeit für eine Sorte mit einem Homogenitätsniveau, das nahe am Populationsstandard liegt.

8.1.3.4 Da die Wahrscheinlichkeit eines Fehlers vom Typ II nicht fest ist, sondern davon abhängt, „wie nichthomogen“ die Kandidatensorte ist, kann diese Wahrscheinlichkeit für verschiedene Nichthomogenitätsgrade berechnet werden. Wie oben erwähnt, gibt dieses Dokument die Wahrscheinlichkeiten eines Fehlers vom Typ II für die Nichthomogenitätsgrade an: 2, 5 und 10 Mal der Populationsstandard.

8.1.3.5 Im allgemeinen nimmt die Wahrscheinlichkeit, Fehler zu begehen, mit der Zunahme der Stichprobengröße ab und mit der Abnahme der Stichprobengröße zu.

8.1.3.6 Für eine gegebene Stichprobengröße kann das Gleichgewicht zwischen den Wahrscheinlichkeiten von Fehlern vom Typ I und vom Typ II durch Änderung der Anzahl zulässiger Abweicher geändert werden.

8.1.3.7 Wenn die Anzahl der zulässigen Abweicher erhöht wird, wird die Wahrscheinlichkeit eines Fehlers vom Typ I reduziert, die Wahrscheinlichkeit eines Fehlers vom Typ II jedoch erhöht. Andererseits wird die Wahrscheinlichkeit eines Fehlers vom Typ I erhöht, wenn die Zahl der zulässigen Abweicher gesenkt wird, während die Wahrscheinlichkeit eines Fehlers vom Typ II reduziert wird.

8.1.3.8 Durch Zulassung einer sehr hohen Anzahl Abweicher ist es möglich, die Wahrscheinlichkeit von Fehlern vom Typ I sehr stark (oder nahezu auf 0) zu senken. Die Wahrscheinlichkeit von Fehlern vom Typ II erhöht sich jedoch nun (unzulässig) stark. Wenn nur eine sehr geringe Anzahl Abweicher zulässig ist, ist das Ergebnis eine geringe Wahrscheinlichkeit von Fehlern vom Typ II und eine (unzulässig) hohe Wahrscheinlichkeit von Fehlern vom Typ I. Der Prozeß des Ausgleichens der Fehler vom Typ I und vom Typ II durch die Wahl der Stichprobengröße und die Anzahl der zulässigen Abweicher wird nachstehend an Beispielen verdeutlicht.

#### 8.1.4 Beispiele

##### Beispiel 1

8.1.4.1 Die Erfahrung zeigt, daß ein angemessener Standard für die betreffende Pflanze 1 % ist. Somit ist der Populationsstandard 1 %. Es wird angenommen, daß eine einzige Prüfung mit höchstens 60 Pflanzen durchgeführt wird. Aus den geeigneten Tabellen (die gewählt wurden, um eine Reihe angestrebter Akzeptanzwahrscheinlichkeiten anzugeben) können folgende Schemata angewandt werden:

Schema	Stichprobengröße	Angestrebte Akzeptanzwahrscheinlichkeit*	Höchstzahl von Abweichern
a	60	90 %	2
b	53	90 %	1
c	60	95 %	2
d	60	99 %	3

8.1.4.2 Folgende Wahrscheinlichkeiten werden für die Fehler vom Typ I und vom Typ II für verschiedene Prozentsätze von Abweichern ermittelt (bezeichnet als  $P_2$ ,  $P_5$  und  $P_{10}$  für das 2-, 5- und 10fache des Populationsstandards):

Schema	Stichprobengröße	Höchstzahl von Abweichern	Fehlerwahrscheinlichkeit (%)			
			Typ I	Typ II		
				$P_2 = 2 \%$	$P_5 = 5 \%$	$P_{10} = 10 \%$
a	60	2	2	88	42	5
b	53	1	10	71	25	3
c	60	2	2	88	42	5
d	60	3	0,3	97	65	14

8.1.4.3 Die Tabelle listet vier verschiedene Schemata auf, und diese sollten untersucht werden, um festzustellen, ob eines davon für die Verwendung geeignet ist. (Die Schemata a und c sind identisch, da es keinen Plan für eine Stichprobengröße von 60 mit einer Wahrscheinlichkeit eines Fehlers vom Typ I zwischen 5 und 10 % gibt). Wird entschieden sicherzustellen, daß die Wahrscheinlichkeit eines Fehlers vom Typ I äußerst gering sein sollte (Schema d), wird die Wahrscheinlichkeit des Fehlers vom Typ II für eine Sorte mit 2, 5 bzw. 10 % Abweichern äußerst hoch (97, 65 und 14 %). Das optimale Gleichgewicht zwischen den Wahrscheinlichkeiten der beiden Fehlertypen scheint durch die Zulassung eines Abweichers in einer Stichprobe von 53 Pflanzen (Schema b) erzielt zu werden.

### Beispiel 2

8.1.4.4 In diesem Beispiel wird eine Pflanze untersucht, bei der der Populationsstandard auf 2 % angesetzt ist und die Zahl der für die Prüfung verfügbaren Pflanzen lediglich 6 beträgt.

8.1.4.5 Unter Verwendung geeigneter Tabellen können folgende Schemata angewandt werden:

\* Siehe Abschnitt 8.1.9.

Schema	Stichproben- größe	Akzeptanz- wahr- schein- lichkeit	Höchstzahl von Abweichern	Fehlerwahrscheinlichkeit (%)			
				Typ I	Typ II		
					P <sub>2</sub> = 4 %	P <sub>5</sub> = 10 %	P <sub>10</sub> = 20 %
a	6	90	1	0,6	98	89	66
b	5	90	0	10	82	59	33
c	6	95	1	0,6	98	89	66
d	6	99	1	0,6	98	89	66
e	6		0	11	78	53	26

8.1.4.6 Schema e der Tabelle wird ermittelt, indem die nachstehend in diesem Dokument angegebenen Formeln (1) und (2) angewandt werden.

8.1.4.7 Dieses Beispiel verdeutlicht die Schwierigkeiten, auf die man stößt, wenn die Stichprobengröße äußerst gering ist. Die Wahrscheinlichkeit, eine nichthomogene Sorte irrtümlicherweise zu akzeptieren (Fehler vom Typ II), ist für alle möglichen Situationen hoch. Selbst wenn alle fünf Pflanzen homogen sein müssen, damit eine Sorte akzeptiert wird (Schema b), beträgt die Wahrscheinlichkeit, eine Sorte mit 20 % Abweichern zu akzeptieren, noch immer 33 %.

8.1.4.8 Es ist anzumerken, daß ein Schema, in dem alle sechs Pflanzen homogen sein müssen (Schema e), etwas geringere Wahrscheinlichkeiten für Fehler vom Typ II ergibt, doch steigt hier die Wahrscheinlichkeit des Fehlers vom Typ I auf 11 %.

8.1.4.9 Schema e kann indessen als die beste Option betrachtet werden, wenn lediglich sechs Pflanzen in einer einzigen Prüfung einer Pflanze, für die der Populationsstandard auf 2 % angesetzt wurde, verfügbar sind.

### Beispiel 3

8.1.4.10 In diesem Beispiel wird erneut die Situation in Beispiel 1 betrachtet, jedoch unter der Annahme, daß die Daten aus zwei Jahren verfügbar sind. Der Populationsstandard beträgt somit 1 % und die Stichprobengröße 120 Pflanzen (60 Pflanzen in jedem der beiden Jahre).

8.1.4.11 Folgende Schemata und Wahrscheinlichkeiten werden aus den geeigneten Tabellen ermittelt:

Schema	Stichproben- größe	Akzeptanz- wahr- schein- lichkeit	Höchstzahl von Abweichern	Fehlerwahrscheinlichkeit (%)			
				Typ I	Typ II		
					P <sub>2</sub> = 2 %	P <sub>5</sub> = 5 %	P <sub>10</sub> = 10 %
a	120	90	3	3	78	15	<0,1
b	110	90	2	10	62	8	<0,1
c	120	95	3	3	78	15	<0,1
d	120	99	4	0,7	91	28	1

8.1.4.12 Hier läßt sich das optimale Gleichgewicht zwischen den Wahrscheinlichkeiten der beiden Fehlertypen mit Schema c erzielen, d. h. nach zwei Jahren insgesamt drei Abweicher unter 120 geprüften Pflanzen zu akzeptieren.

8.1.4.13 Im anderen Falle läßt sich ein zweiphasiges Prüfungsverfahren aufstellen. Ein solches Verfahren läßt sich für diesen Fall ermitteln, indem die nachstehend in diesem Dokument angeführten Formeln (3) und (4) angewandt werden.

8.1.4.14 Folgende Schemata lassen sich ermitteln:

Schema	Stichproben- größe	Akzeptanz- wahr- schein- lichkeit	Größe Zahl für die Akzeptanz nach Jahr 1	Größe Zahl vor der Zurückweisung im Jahr 1	Größe nach 2 Jahren zu akzeptierende Zahl
e	60	90	kann nie akzeptiert werden	2	3
f	60	95	kann nie akzeptiert werden	2	3
g	60	99	kann nie akzeptiert werden	3	4
h	58	90	1	2	2

8.1.4.15 Unter Anwendung der Formeln (3), (4) und (5) lassen sich folgende Fehlerwahrscheinlichkeiten ermitteln:

Schema	Fehlerwahrscheinlichkeit (%)				Wahrscheinlichkeit der Prüfung in einem 2. Jahr
	Typ I	Typ II			
		$P_2 = 2\%$	$P_5 = 5\%$	$P_{10} = 10\%$	
e	4	75	13	0,1	100
f	4	75	13	0,1	100
g	1	90	27	0,5	100
h	10	62	9	0,3	36

8.1.4.16 Die Schemata e und f ergeben beide eine Wahrscheinlichkeit von 4 % für die Zurückweisung einer homogenen Sorte (Fehler vom Typ I) und eine Wahrscheinlichkeit von 13 % für die Akzeptanz einer Sorte mit 5 % Abweichern (Fehler vom Typ II). Die Entscheidung lautet:

- die Sorte nie nach einem Jahr akzeptieren
- mehr als 2 Abweicher im Jahr 1: die Sorte zurückweisen und die Prüfung abbrechen
- 0 bis und mit 2 Abweicher im Jahr 1: ein zweites Prüfungsjahr einlegen
- höchstens 3 Abweicher nach 2 Jahren: die Sorte akzeptieren
- mehr als 3 Abweicher nach 2 Jahren: die Sorte zurückweisen

8.1.4.17 Im anderen Falle kann Schema a oder h gewählt werden, doch scheint Schema g eine zu hohe Wahrscheinlichkeit von Fehlern vom Typ II im Vergleich zur Wahrscheinlichkeit von Fehlern vom Typ I aufzuweisen. Es besteht beispielsweise eine Wahrscheinlichkeit von 1 %, eine homogene Sorte zurückzuweisen (Fehler vom Typ I), und eine Wahrscheinlichkeit von 27 %, eine Sorte mit 5 % Abweichern zu akzeptieren (Fehler vom Typ II).

8.1.4.18 Schema h hat den Vorzug, häufig eine endgültige Entscheidung nach der ersten Prüfung (Prüfungsjahr) zuzulassen, doch ist infolgedessen die Wahrscheinlichkeit eines Fehlers vom Typ I höher. In diesem Fall ist eine Wahrscheinlichkeit von 10 % vorhanden, daß eine homogene Sorte zurückgewiesen wird (Fehler vom Typ I) und eine Wahrscheinlichkeit von 9 %, daß eine Sorte mit 5 % Abweichern akzeptiert wird (Fehler vom Typ II).

#### Beispiel 4

8.1.4.19 In diesem Beispiel wird angenommen, daß der Populationsstandard 3 % beträgt und daß in jedem der beiden Prüfungsjahre je acht Pflanzen verfügbar sind.

8.1.4.20 Aus den geeigneten Tabellen geht hervor:

Schema	Stichproben- größe	Akzeptanz- wahr- schein- lichkeit	Höchstzahl von Abweichern	Fehlerwahrscheinlichkeit ( %)			
				Typ I	Typ II		
					P <sub>2</sub> = 6 %	P <sub>5</sub> = 15 %	P <sub>10</sub> = 30 %
a	16	90	1	8	78	28	3
b	16	95	2	1	93	56	10
c	16	99	3	0,1	99	79	25

8.1.4.21 Hier läßt sich das optimale Gleichgewicht zwischen den beiden Fehlertypen mit Schema a erzielen.

8.1.4.22 Das Verfahren „seedcalc“ der Internationalen Vereinigung für Saatgutprüfung (ISTA) kann für die Berechnung der Fehler vom Typ I und vom Typ II angewandt werden. „Seedcalc“ ist unter folgender Website-Adresse verfügbar:  
[http://www.seedtest.org/en/stats\\_tool\\_box\\_content---1--1143.html](http://www.seedtest.org/en/stats_tool_box_content---1--1143.html)

### 8.1.5 Einführung in die Tabellen und Abbildungen

8.1.5.1 Im Abschnitt TABELLEN UND ABBILDUNGEN (Teil II Abschnitt 8.1.10) sind 7 Paare von Tabellen und Abbildungen enthalten, die verschiedenen Kombinationen des Populationsstandards und der Akzeptanzwahrscheinlichkeit entsprechen. Diese sind für die Anwendung auf eine einzige Abweicherprüfung vorgesehen. Eine Übersicht über die Tabellen und Abbildungen ist in Tabelle A enthalten.

8.1.5.2 Jede Tabelle zeigt die Höchstzahlen von Abweichern (k) mit den entsprechenden Spannweiten der Stichprobengrößen (n) für den gegebenen Populationsstandard und die gegebene Akzeptanzwahrscheinlichkeit. In Tabelle 1 (Populationsstandard 10%, Akzeptanzwahrscheinlichkeit 95%) beispielsweise liegt die entsprechende Stichprobengröße n für eine auf 2 festgelegte Höchstzahl von Abweichern in der Spannweite 11 bis 22. Gleichmaßen sollte, wenn die Höchstzahl der Abweicher (k) 10 beträgt, die entsprechende anzuwendende Stichprobengröße (n) in der Spannweite 126 bis 141 liegen.

8.1.5.3 Für kleine Stichprobengrößen werden dieselben Informationen in den entsprechenden Abbildungen (1 bis 7) graphisch dargestellt. Diese zeigen das tatsächliche Risiko der Zurückweisung einer homogenen Sorte und die Wahrscheinlichkeit der Akzeptanz einer Sorte mit einem wahren Anteil von Abweichern, der das 2fache (2P), 5fache (5P) und 10fache (10P) des Populationsstandards beträgt. (Um die Lektüre der Abbildung zu erleichtern, werden die Risiken für die individuellen Stichprobengrößen durch Linien verbunden, obwohl die Wahrscheinlichkeit nur für jede individuelle Stichprobengröße berechnet werden kann.)

Tabelle A. Übersicht über Tabellen und Abbildungen 1 bis 7.

Populationsstandard %	Akzeptanzwahrscheinlichkeit %	Siehe Tabelle und Abbildung Nr.
10	>95	1
5	>95	2
3	>95	3
2	>95	4
1	>95	5
0,5	>95	6
0,1	>95	7

8.1.5.4 Zur Verwendung der Tabellen wird folgendes Verfahren vorgeschlagen:

- a) den entsprechenden Populationsstandard wählen.
- b) das Entscheidungsschema mit dem besten Gleichgewicht zwischen den Fehlerwahrscheinlichkeiten wählen.

8.1.5.5 Die Verwendung der Tabellen und Abbildungen wird im Abschnitt mit den Beispielen veranschaulicht.

#### 8.1.6 Verfahren für eine einzige Prüfung

Die mathematischen Berechnungen beruhen auf der Binomialverteilung, und in der Regel werden folgende Begriffe verwendet:

a) Der in einem spezifischen Fall zu akzeptierende Prozentsatz von Abweichern wird als „Populationsstandard“ bezeichnet und mit dem Buchstaben P dargestellt.

b) Die „Akzeptanzwahrscheinlichkeit“ ist die Wahrscheinlichkeit, eine Sorte mit P % Abweichern zu akzeptieren. Da die Zahl der Abweicher diskret ist, variiert jedoch die tatsächliche Wahrscheinlichkeit der Akzeptanz einer homogenen Sorte mit der Stichprobengröße, sie ist jedoch stets größer als oder gleich wie die „Akzeptanzwahrscheinlichkeit“. Die Akzeptanzwahrscheinlichkeit wird in der Regel mit  $100 - \alpha$  bezeichnet, wobei  $\alpha$  die Wahrscheinlichkeit der Zurückweisung einer Sorte mit P % Abweichern ist (d. h. die Wahrscheinlichkeit eines Fehlers vom Typ I). In der Praxis haben zahlreiche Sorten weniger als P % Abweicher; deshalb ist der Fehler vom Typ I für diese Sorten tatsächlich kleiner als  $\alpha$ .

c) Die Anzahl geprüfter Pflanzen in einer Zufallsstichprobe wird Stichprobengröße genannt und mit n bezeichnet.

d) Die Höchstzahl der tolerierten Abweicher in einer Zufallsstichprobe der Größe n wird als k bezeichnet.

e) Die Wahrscheinlichkeit, eine Sorte mit mehr als P % Abweichern, z. B. mit  $P_q$  % Abweichern, zu akzeptieren, wird mit dem Buchstaben  $\beta$  oder mit  $\beta_q$  bezeichnet.

f) die mathematischen Formeln für die Berechnung der Wahrscheinlichkeiten sind:

$$\alpha = 100 - 100 \sum_{i=0}^k \binom{n}{i} P^i (1-P)^{n-i} \quad (1)$$

$$\beta_q = 100 \sum_{i=0}^k \binom{n}{i} P_q^i (1-P_q)^{n-i} \quad (2)$$

P und  $P_q$  sind hier als Proportionen ausgedrückt, d. h. Prozente dividiert durch 100.

### 8.1.7 Methode für mehr als eine einzige Prüfung (Prüfungsjahr)

#### *8.1.7.1 Einleitung*

8.1.7.1.1 Häufig wird eine Kandidatensorte während zwei (oder drei) Jahren angebaut. Dann taucht die Frage auf, wie die Informationen über die Homogenität aus den einzelnen Jahren zu kombinieren sind. Zwei Methoden werden geschildert:

a) Die Entscheidung nach zwei (oder drei) Jahren treffen, beruhend auf der Gesamtzahl der geprüften Pflanzen und der Gesamtzahl der erfaßten Abweicher (kombinierte Prüfung).

b) Das Ergebnis des ersten Jahres verwenden, um festzustellen, ob die Daten eine klare Entscheidung andeuten (Zurückweisung oder Akzeptanz). Ist die Entscheidung nicht klar, mit dem zweiten Jahr fortfahren und nach dem zweiten Jahr entscheiden (Zweiphasenprüfung).

8.1.7.1.2 Es gibt indessen einige Alternativen (z. B. kann in jedem Jahr eine Entscheidung getroffen werden, und eine endgültige Entscheidung kann durch die Zurückweisung der Kandidatensorte getroffen werden, wenn diese zu viele Abweicher in beiden (oder in zwei von drei Jahren) aufweist). Ferner gibt es Komplikationen, wenn eine Prüfung in mehr als einem einzigen Jahr durchgeführt wird. Deshalb wird vorgeschlagen, einen Statistiker hinzuziehen, wenn eine Prüfung von zwei (oder mehr) Jahren durchzuführen ist.

#### *8.1.7.2 Kombinierte Prüfung*

Die Stichprobengröße bei Prüfung i ist  $n_i$ . Nach der letzten Prüfung haben wir somit die gesamte Stichprobengröße  $n = \sum n_i$ . Ein Entscheidungsschema wird auf genau dieselbe Weise aufgestellt, als ob diese gesamte Stichprobengröße in einer einzigen Prüfung ermittelt worden wäre. So wird die Gesamtzahl der während der Prüfungen erfaßten Abweicher mit der Höchstzahl der vom gewählten Entscheidungsschema zugelassenen Abweicher verglichen.

#### *8.1.7.3 Zweiphasenprüfung*

8.1.7.3.1 Die Methode für eine Zweijahresprüfung läßt sich folgendermaßen beschreiben: Im ersten Jahr eine Stichprobe der Größe n nehmen. Die Kandidatensorte zurückweisen, wenn

mehr als  $r_1$  Abweicher erfaßt werden, und die Kandidatensorte annehmen, wenn weniger als  $a_1$  Abweicher erfaßt werden. Ansonsten zum zweiten Jahr übergehen und eine Stichprobe der Größe  $n$  (wie im ersten Jahr) nehmen und die Kandidatensorte zurückweisen, wenn die Gesamtzahl der in der Zweijahresprüfung erfaßten Abweicher größer als  $r$  ist. Andernfalls die Kandidatensorte akzeptieren. Die endgültigen Risiken und die erwartete Stichprobengröße bei

$$\begin{aligned} \alpha &= P(K_1 > r_1) + P(K_1 + K_2 > r \mid K_1) \\ &= P(K_1 > r_1) + P(K_2 > r - K_1 \mid K_1) \\ &= \sum_{i=r_1+1}^n \binom{n}{i} P^i (1-P)^{n-i} + \sum_{i=a_1}^{r_1} \binom{n}{i} P^i (1-P)^{n-i} \sum_{j=r-i+1}^n \binom{n}{j} P^j (1-P)^{n-j} \quad (3) \end{aligned}$$

$$\begin{aligned} \beta_q &= P(K_1 < a_1) + P(K_1 + K_2 \leq r \mid K_1) \\ &= P(K_1 < a_1) + P(K_2 \leq r - K_1 \mid K_1) \\ &= \sum_{i=0}^{a_1-1} \binom{n}{i} P_q^i (1-P_q)^{n-i} + \sum_{i=a_1}^{r_1} \binom{n}{i} P_q^i (1-P_q)^{n-i} \sum_{j=0}^{r-i} \binom{n}{j} P_q^j (1-P_q)^{n-j} \quad (4) \end{aligned}$$

einem derartigen Verfahren lassen sich wie folgt berechnen:

$$n_e = n \left( 1 + \sum_{i=a_1}^{r_1} \binom{n}{i} P^i (1-P)^{n-i} \right) \quad (5)$$

wobei

$P$  = Populationsstandard

$\alpha$  = Wahrscheinlichkeit eines tatsächlichen Fehlers vom Typ I für  $P$

$\beta_q$  = Wahrscheinlichkeit eines tatsächlichen Fehlers vom Typ II für  $q P$

$n_e$  = erwartete Stichprobengröße

$r_1, a_1$  und  $r$  Entscheidungsparameter sind

$P_q$  = das  $q$ -fache des Populationsstandards =  $q P$

$K_1$  und  $K_2$  die Zahl der im Jahr 1 bzw. im Jahr 2 festgestellten Abweicher sind.

Die Entscheidungsparameter  $a_1, r_1$  und  $r$  können gemäß folgenden Kriterien gewählt werden:

- $\alpha$  muß weniger als  $\alpha_0$  sein, wobei  $\alpha_0$  der maximale Fehler vom Typ I ist, d. h.  $\alpha_0$  ist 100 minus die erforderliche Akzeptanzwahrscheinlichkeit
- $\beta_q$  (für  $q=5$ ) sollte möglichst gering, jedoch nicht kleiner als  $\alpha_0$  sein
- wenn  $\beta_q$  (für  $q=5$ )  $< \alpha_0$ , sollte  $n_e$  möglichst gering sein.

8.1.7.3.2 Es stehen allerdings auch andere Strategien zur Verfügung, und es werden hier keine Tabellen/Abbildungen gezeigt, da es mehrere verschiedene Entscheidungsschemata geben könnte, die einer bestimmten Risikoserie entsprechen. Es wird vorgeschlagen, einen

Statistiker zu konsultieren, wenn eine zweiphasige Prüfung (oder andere Sequenzprüfungen) erforderlich sind.

#### 8.1.7.4 Sequenzprüfungen

Die obenerwähnte zweiphasige Prüfung ist eine Art Sequenzprüfung, bei der das Ergebnis der ersten Phase bestimmt, ob die Prüfung während einer zweiten Phase fortgesetzt werden muß. Es können auch andere Arten von Sequenzprüfungen angewandt werden. Es kann sachdienlich sein, solche Prüfungen in Betracht zu ziehen, wenn die praktische Arbeit erlaubt, daß in bestimmten Phasen der Prüfung Analysen von Abweichern vorgenommen werden. Die Entscheidungsschemata für derartige Methoden lassen sich auf unterschiedliche Arten aufstellen, und es wird vorgeschlagen, einen Statistiker hinzuzuziehen, wenn Sequenzverfahren anzuwenden sind.

#### 8.1.8 Anmerkung zum Ausgleich der Fehler vom Typ I und vom Typ II

8.1.8.1 Da die Zahl der Abweicher diskret ist, können wir in der Regel keine Fehler vom Typ I erzielen, die schöne, präselektionierte Zahlen sind. Das Schema a im obigen Beispiel 2 mit 6 Pflanzen zeigte, daß wir kein  $\alpha$  von 10 % erzielen konnten – unser tatsächliches  $\alpha$  betrug 0,6 %. Eine Änderung der Stichprobengröße hat schwankende  $\alpha$ - und  $\beta$ -Werte zur Folge. Abbildung 3 – als Beispiel – zeigt, daß sich  $\alpha$  bei bestimmten Stichprobengrößen seinen Nominalwerten stärker annähert, und dies ist auch die Stichprobengröße, bei der  $\beta$  verhältnismäßig gering ist.

8.1.8.2 Im allgemeinen sind große Stichprobengrößen von Vorteil. Bei gleicher Akzeptanzwahrscheinlichkeit neigt eine größere Stichprobe dazu, eine proportional geringere Wahrscheinlichkeit von Fehlern vom Typ II zu haben. Kleine Stichprobengrößen führen zu einer hohen Wahrscheinlichkeit, nichthomogene Sorten zu akzeptieren. Deshalb sollte die Stichprobengröße so gewählt werden, daß sich ein annehmbar geringes Niveau von Fehlern vom Typ II ergibt. Geringfügige Erhöhungen der Stichprobengröße sind jedoch nicht immer von Vorteil. Eine Stichprobengröße von fünf ergibt beispielsweise  $\alpha = 10\%$  und  $\beta_2 = 82\%$ , während eine Stichprobengröße von sechs  $\alpha = 0,6\%$  und  $\beta_2 = 98\%$  ergibt. Es scheint, daß die Stichprobengrößen, die  $\alpha$ -Werte ergeben, die der Akzeptanzwahrscheinlichkeit recht genau entsprechen, die größten in einer Reihe von Stichprobengrößen mit einer genau angegebenen Höchstzahl von Abweichern sind. So sollten die größten Stichprobengrößen in der Spannweite von Stichprobengrößen mit einer gegebenen Höchstzahl von Abweichern verwendet werden.

#### 8.1.9 Definition der statistischen Begriffe und Symbole

Die verwendeten statistischen Begriffe und Symbole sind wie folgt definiert:

*Populationsstandard.* Der Prozentsatz der zulässigen Abweicher, wenn alle Individuen einer Sorte geprüft werden könnten. Der Populationsstandard wird für die betreffende Pflanze festgesetzt und beruht auf der Erfahrung.

*Akzeptanzwahrscheinlichkeit.* Die Wahrscheinlichkeit, daß eine Sorte mit P % Abweichern akzeptiert wird. P ist hier der Populationsstandard. Es ist jedoch anzumerken, daß die tatsächliche Wahrscheinlichkeit, eine homogene Sorte zu akzeptieren, stets größer als oder gleich wie die Akzeptanzwahrscheinlichkeit in den Überschriften der Tabellen und

Abbildungen ist. Die Wahrscheinlichkeit, eine homogene Sorte zu akzeptieren, und die Wahrscheinlichkeit eines Fehlers vom Typ I belaufen sich auf 100 %. Wenn der Fehler vom Typ I z. B. 4 % beträgt, dann beträgt die Wahrscheinlichkeit, eine homogene Sorte zu akzeptieren,  $100 - 4 = 96$  % (siehe z. B. Abbildung I für  $n=50$ ). Der Fehler vom Typ I ist im Diagramm in den Abbildungen durch die sägeartigen Spitzen zwischen null und der oberen Grenze des Fehlers vom Typ I (z. B. 10 in Abbildung 1) angegeben. Die Entscheidungsschemata werden so definiert, daß die tatsächliche Wahrscheinlichkeit, daß eine homogene Sorte akzeptiert wird, stets größer als oder gleich wie die Akzeptanzwahrscheinlichkeit in der Überschrift der Tabelle ist.

*Fehler vom Typ I.* Der Fehler, eine homogene Sorte zurückzuweisen.

*Fehler vom Typ II.* Der Fehler, eine zu stark nichthomogene Sorte zu akzeptieren.

P Populationsstandard

$P_q$  Der angenommene wahre Prozentsatz von Abweichern in einer nichthomogenen Sorte.  
 $P_q = q P$ .

Im diesem Dokument beträgt  $q$  2, 5 oder 10. Dies sind nur drei Beispiele zur Darstellung der Fehler vom Typ II. Der tatsächliche Prozentsatz von Abweichern in einer Sorte kann jeden Wert annehmen. Wir können z. B. verschiedene Sorten prüfen, die faktisch 1,6 %, 3,8 %, 0,2 %, ... Abweicher aufweisen.

$n$  Stichprobengröße

$k$  Höchstzahl der zulässigen Abweicher

$\alpha$  Wahrscheinlichkeit eines Fehlers vom Typ I

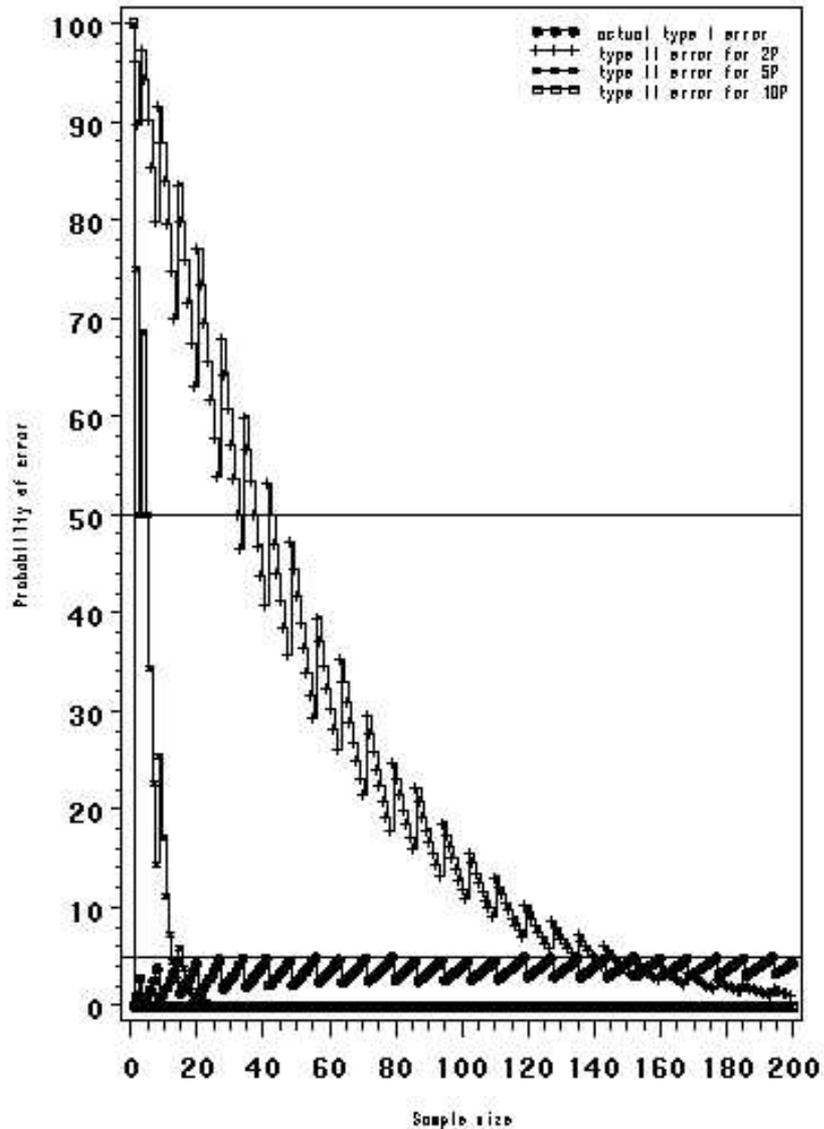
$\beta$  Wahrscheinlichkeit eines Fehlers vom Typ II

8.1.10 Tabellen und Abbildungen

Tabelle und Abbildung 1:

Populationsstandard = 10 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95\%$   
 n = Stichprobengröße k = Höchstzahl von Abweichern

	n		k
	1	bis	3
	4	bis	8
	9	bis	14
	15	bis	20
	21	bis	27
	28	bis	34
	35	bis	41
	42	bis	48
	49	bis	56
	57	bis	63
	64	bis	71
	72	bis	79
	80	bis	86
	87	bis	94
	95	bis	102
	103	bis	110
	111	bis	119
	120	bis	127
	128	bis	135
	136	bis	143
	144	bis	152
	153	bis	160
	161	bis	168
	169	bis	177
	178	bis	185
	186	bis	194
	195	bis	200



Fehlerwahrscheinlichkeit

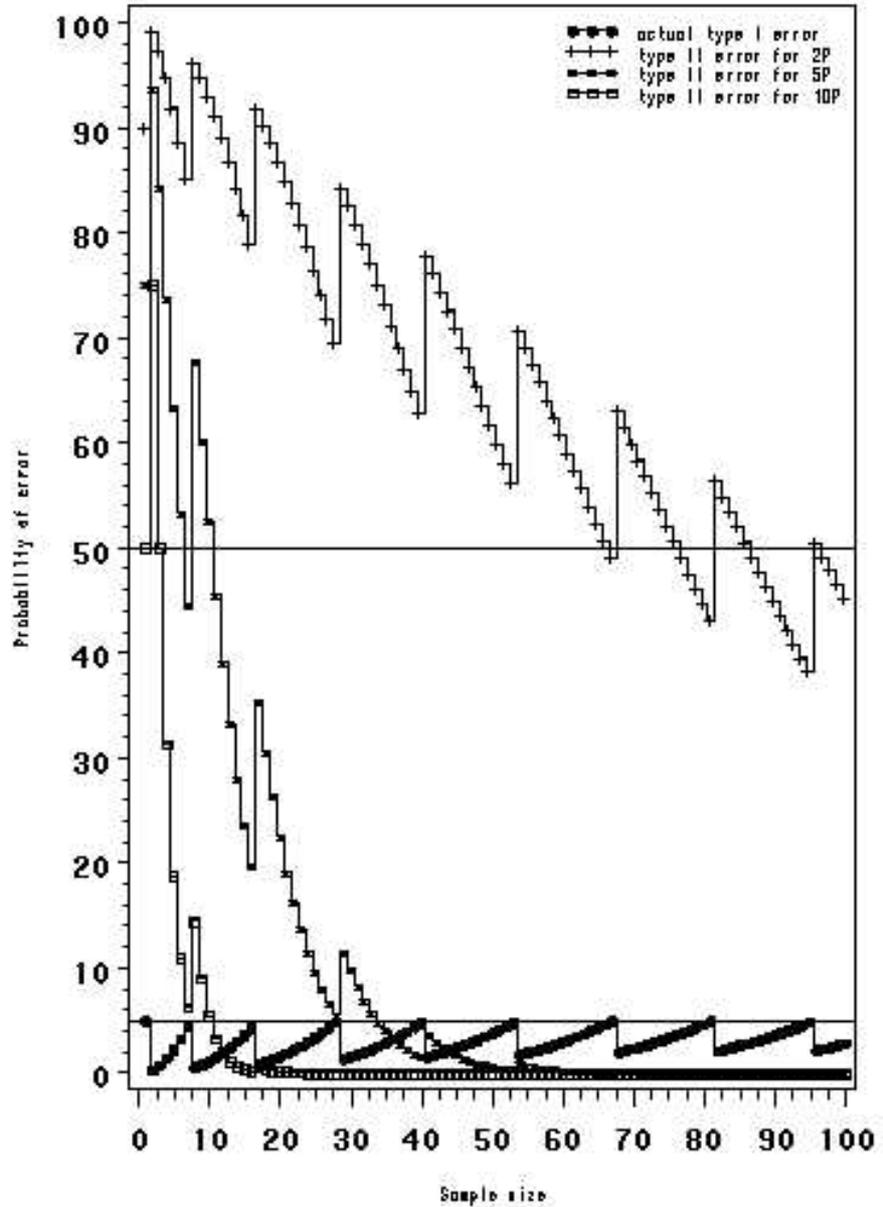
Stichprobengröße

tatsächlicher Fehler vom Typ I  
 Fehler vom Typ II für 2P  
 Fehler vom Typ II für 5P  
 Fehler vom Typ II für 10P

Tabelle und Abbildung 2:

Populationsstandard = 5 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95$  %  
 n = Stichprobengröße k = Höchstzahl von Abweichern

	n	k
1	bis	1
2	bis	7
8	bis	16
17	bis	28
29	bis	40
41	bis	53
54	bis	67
68	bis	81
82	bis	95
96	bis	110
111	bis	125
126	bis	140
141	bis	155
156	bis	171
172	bis	187
188	bis	203
204	bis	219
220	bis	235
236	bis	251
252	bis	268
269	bis	284
285	bis	300
301	bis	317
318	bis	334
335	bis	351
352	bis	367
368	bis	384
385	bis	401
402	bis	418
419	bis	435
436	bis	452
453	bis	469
470	bis	487
488	bis	504
505	bis	521
522	bis	538
539	bis	556
557	bis	573
574	bis	590
591	bis	608
609	bis	625
626	bis	643
644	bis	660
661	bis	678
679	bis	696
697	bis	713
714	bis	731
732	bis	748
749	bis	766
767	bis	784
785	bis	802
803	bis	819
820	bis	837
838	bis	855
856	bis	873
874	bis	891
892	bis	909
910	bis	926
927	bis	944
945	bis	962
963	bis	980
981	bis	998

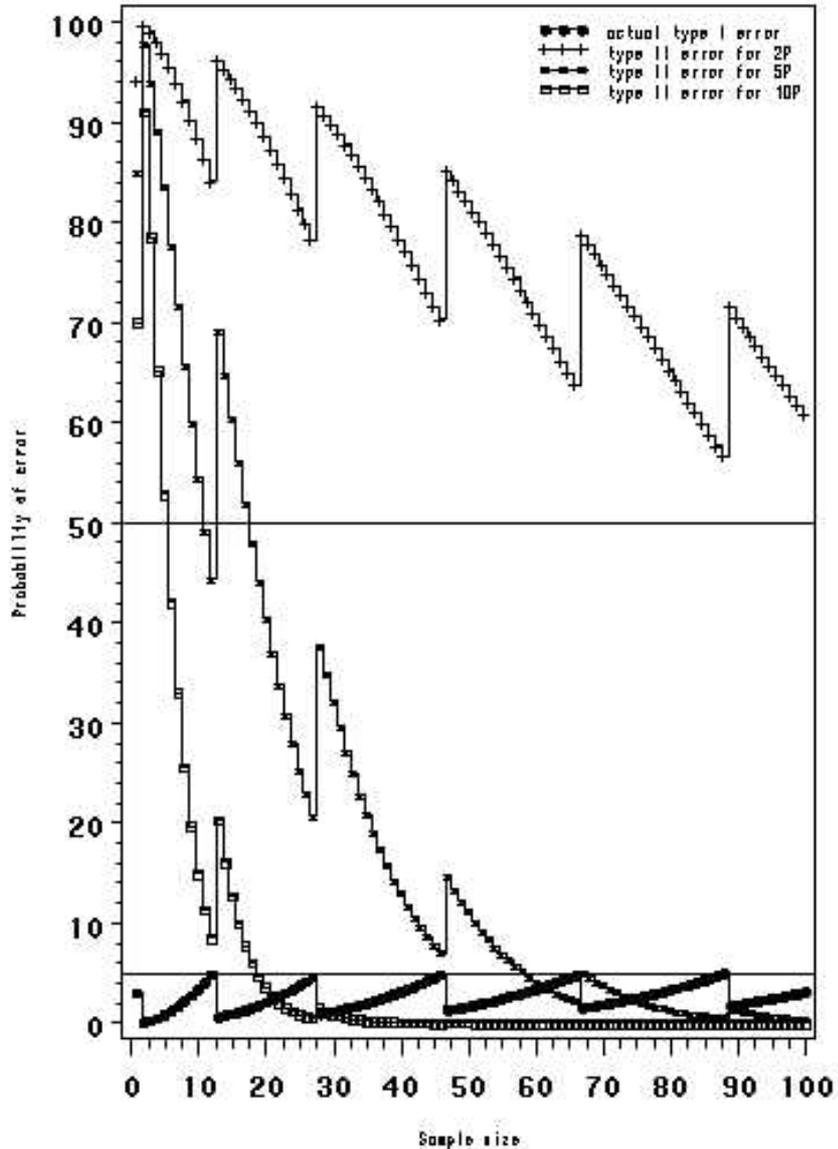


Fehlerwahrscheinlichkeit      Stichprobengröße      tatsächlicher Fehler vom Typ I  
 Fehler vom Typ II für 2P  
 Fehler vom Typ II für 5P  
 Fehler vom Typ II für 10P

Tabelle und Abbildung 3:

Populationsstandard = 3 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95$  %  
 n = Stichprobengröße k = Höchstzahl von Abweichern

	n	k
	1 bis	1
	2 bis	12
	13 bis	27
	28 bis	46
	47 bis	66
	67 bis	88
	89 bis	110
	111 bis	134
	135 bis	158
	159 bis	182
	183 bis	207
	208 bis	232
	233 bis	258
	259 bis	284
	285 bis	310
	311 bis	337
	338 bis	363
	364 bis	390
	391 bis	417
	418 bis	444
	445 bis	472
	473 bis	499
	500 bis	527
	528 bis	554
	555 bis	582
	583 bis	610
	611 bis	638
	639 bis	666
	667 bis	695
	696 bis	723
	724 bis	751
	752 bis	780
	781 bis	809
	810 bis	837
	838 bis	866
	867 bis	895
	896 bis	924
	925 bis	952
	953 bis	981
	982 bis	1010
	1011 bis	1040
	1041 bis	1069
	1070 bis	1098
	1099 bis	1127
	1128 bis	1156
	1157 bis	1186
	1187 bis	1215
	1216 bis	1244
	1245 bis	1274
	1275 bis	1303
	1304 bis	1333
	1334 bis	1362
	1363 bis	1392
	1393 bis	1422
	1423 bis	1451
	1452 bis	1481
	1482 bis	1511
	1512 bis	1541
	1542 bis	1570
	1571 bis	1600
	1601 bis	1630
	1631 bis	1660
	1661 bis	1690
	1691 bis	1720
	1721 bis	1750
	1751 bis	1780
	1781 bis	1810
	1811 bis	1840
	1841 bis	1870
	1871 bis	1900

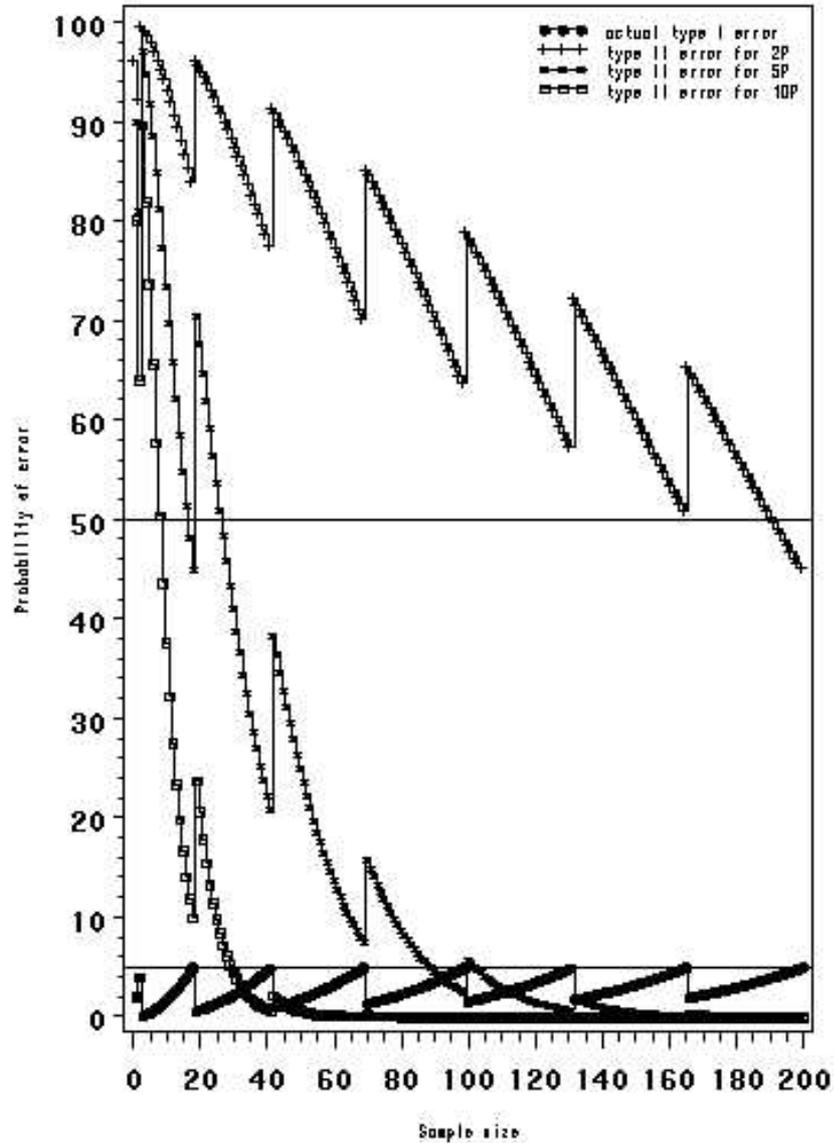


Fehlerwahrscheinlichkeit      Stichprobengröße      tatsächlicher Fehler vom Typ I  
 Fehler vom Typ II für 2P  
 Fehler vom Typ II für 5P  
 Fehler vom Typ II für 10P

Tabelle und Abbildung 4:

Populationsstandard = 2 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95$  %  
 n = Stichprobengröße k = Höchstzahl von Abweichern

n	k
1 bis 2	0
3 bis 18	1
19 bis 41	2
42 bis 69	3
70 bis 99	4
100 bis 131	5
132 bis 165	6
166 bis 200	7
201 bis 236	8
237 bis 273	9
274 bis 310	10
311 bis 348	11
349 bis 386	12
387 bis 425	13
426 bis 464	14
465 bis 504	15
505 bis 544	16
545 bis 584	17
585 bis 624	18
625 bis 665	19
666 bis 706	20
707 bis 747	21
748 bis 789	22
790 bis 830	23
831 bis 872	24
873 bis 914	25
915 bis 956	26
957 bis 998	27
999 bis 1040	28
1041 bis 1083	29
1084 bis 1126	30
1127 bis 1168	31
1169 bis 1211	32
1212 bis 1254	33
1255 bis 1297	34
1298 bis 1340	35
1341 bis 1383	36
1384 bis 1427	37
1428 bis 1470	38
1471 bis 1514	39
1515 bis 1557	40
1558 bis 1601	41
1602 bis 1645	42
1646 bis 1689	43
1690 bis 1732	44
1733 bis 1776	45
1777 bis 1820	46
1821 bis 1864	47
1865 bis 1909	48
1910 bis 1953	49
1954 bis 1997	50
1998 bis 2000	51

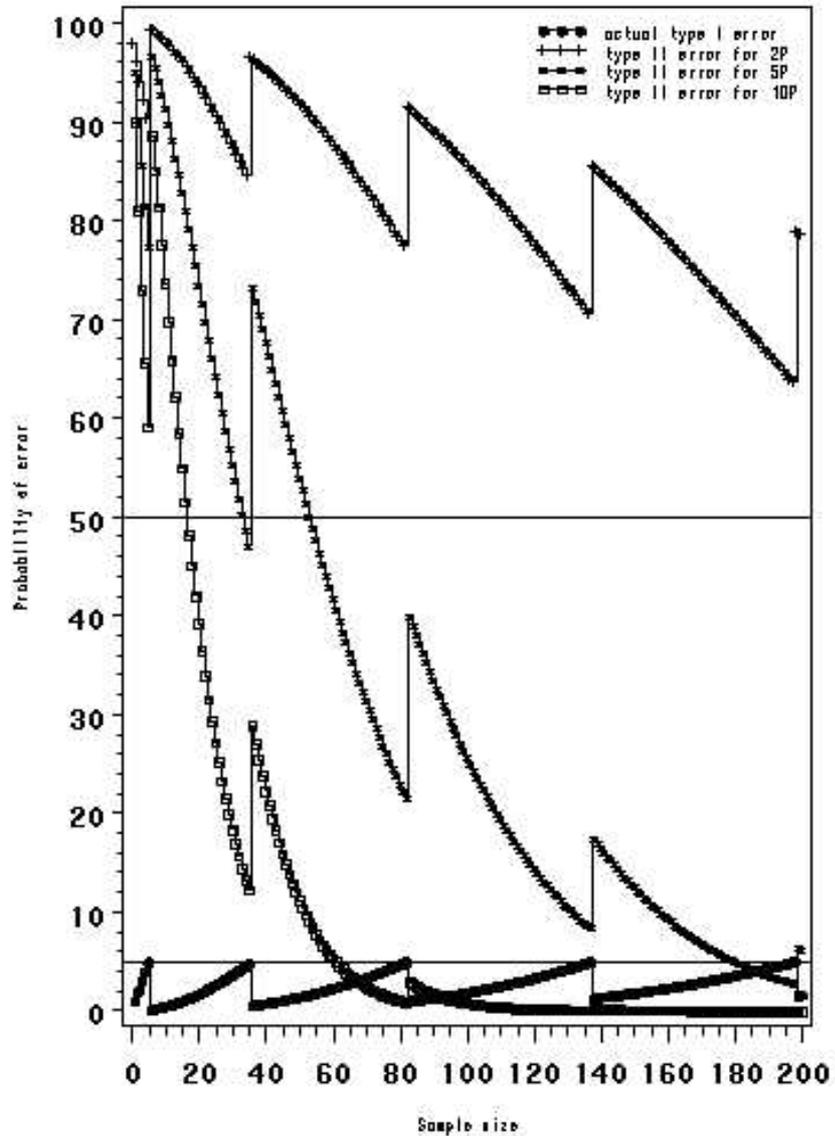


Fehlerwahrscheinlichkeit	Stichprobengröße	tatsächlicher Fehler vom Typ I
		Fehler vom Typ II für 2P
		Fehler vom Typ II für 5P
		Fehler vom Typ II für 10P

Tabelle und Abbildung 5:

Populationsstandard = 1 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95\%$   
 n = Stichprobengröße k = Höchstzahl von Abweichern

n	k
1 bis 5	0
6 bis 35	1
36 bis 82	2
83 bis 137	3
138 bis 198	4
199 bis 262	5
263 bis 329	6
330 bis 399	7
400 bis 471	8
472 bis 544	9
545 bis 618	10
619 bis 694	11
695 bis 771	12
772 bis 848	13
849 bis 927	14
928 bis 1006	15
1007 bis 1085	16
1086 bis 1166	17
1167 bis 1246	18
1247 bis 1328	19
1329 bis 1410	20
1411 bis 1492	21
1493 bis 1575	22
1576 bis 1658	23
1659 bis 1741	24
1742 bis 1825	25
1826 bis 1909	26
1910 bis 1993	27
1994 bis 2078	28
2079 bis 2163	29
2164 bis 2248	30
2249 bis 2333	31
2334 bis 2419	32
2420 bis 2505	33
2506 bis 2591	34
2592 bis 2677	35
2678 bis 2763	36
2764 bis 2850	37
2851 bis 2937	38
2938 bis 3000	39

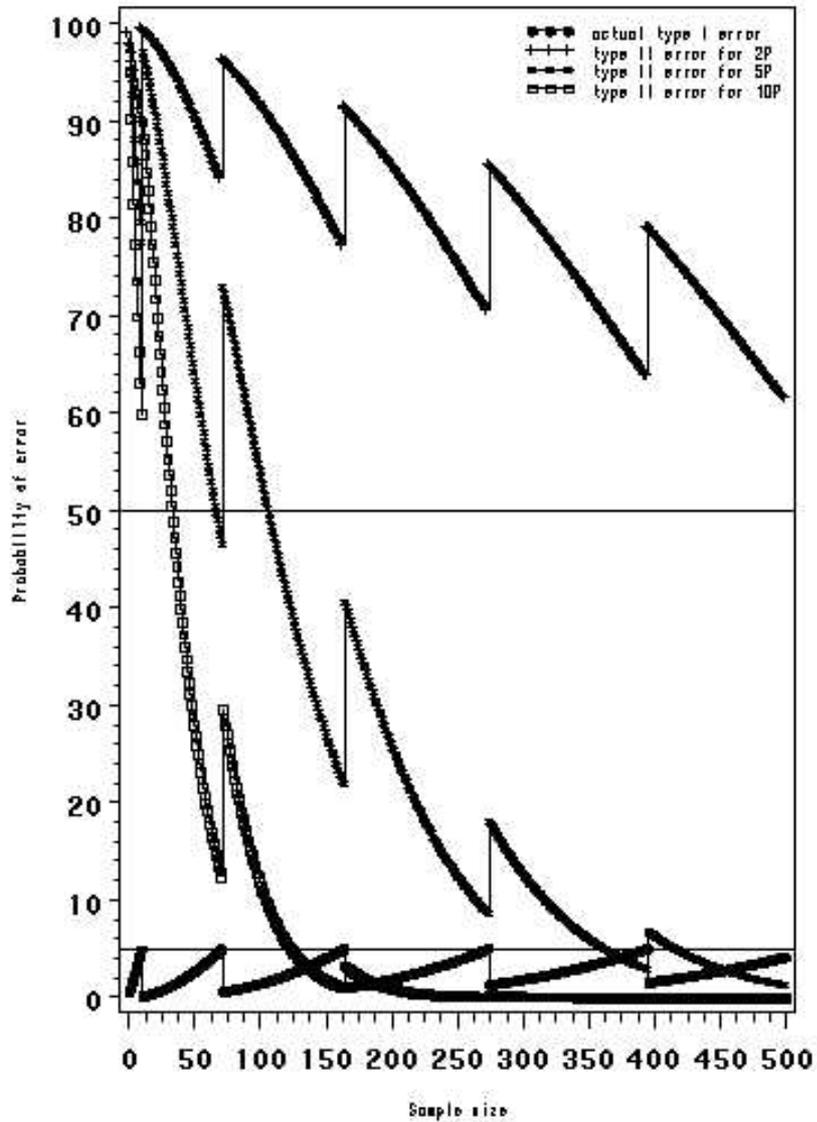


Fehlerwahrscheinlichkeit      Stichprobengröße      tatsächlicher Fehler vom Typ I  
 Fehler vom Typ II für 2P  
 Fehler vom Typ II für 5P  
 Fehler vom Typ II für 10P

Tabelle und Abbildung 6:

Populationsstandard = 5 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95$  %  
 n = Stichprobengröße k = Höchstzahl von Abweichern

n	k
1 bis 10	0
11 bis 71	1
72 bis 164	2
165 bis 274	3
275 bis 395	4
396 bis 523	5
524 bis 658	6
659 bis 797	7
798 bis 940	8
941 bis 1086	9
1087 bis 1235	10
1236 bis 1386	11
1387 bis 1540	12
1541 bis 1695	13
1696 bis 1851	14
1852 bis 2009	15
2010 bis 2169	16
2170 bis 2329	17
2330 bis 2491	18
2492 bis 2653	19
2654 bis 2817	20
2818 bis 2981	21
2982 bis 3000	22

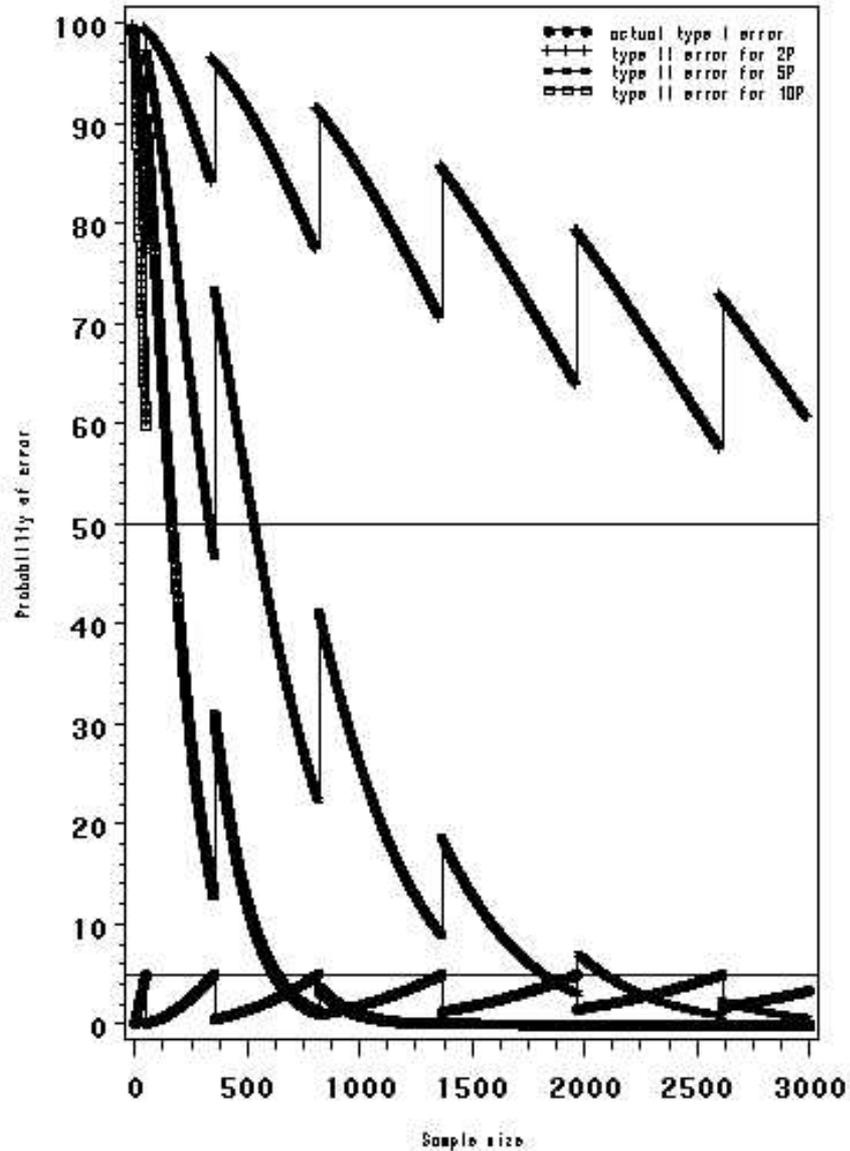


Fehlerwahrscheinlichkeit	Stichprobengröße	tatsächlicher Fehler vom Typ I
		Fehler vom Typ II für 2P
		Fehler vom Typ II für 5P
		Fehler vom Typ II für 10P

Tabelle und Abbildung 7:

Populationsstandard = 1 %  
 Akzeptanzwahrscheinlichkeit  $\geq 95\%$   
 n = Stichprobengröße k = Höchstzahl von Abweichern

n	k
1 bis 51	0
52 bis 355	1
356 bis 818	2
819 bis 1367	3
1368 bis 1971	4
1972 bis 2614	5
2615 bis 3000	6



Fehlerwahrscheinlichkeit	Stichprobengröße	tatsächlicher Fehler vom Typ I
		Fehler vom Typ II für 2P
		Fehler vom Typ II für 5P
		Fehler vom Typ II für 10P

## **9. DAS KOMBINIERTES HOMOGENITÄTSKRITERIUM ÜBER MEHRERE JAHRE (COYU)**

### **9.1 Zusammenfassung der Voraussetzungen für die Anwendung des Verfahrens**

- für quantitative Merkmale;
- wenn die Beobachtungen auf Einzelpflanzenbasis über zwei oder mehrere Jahre erfolgen.
- wenn es Unterschiede zwischen Pflanzen einer Sorte gibt, die vielmehr eine quantitative Variation als das Vorhandensein von Abweichern darstellen;
- Es wird empfohlen, daß es mindestens 20 Freiheitsgrade für die Schätzung der Varianz für die vergleichbaren Sorten in der COYU-Analyse geben sollte.

Vergleichbare Sorten sind Sorten desselben Typs innerhalb derselben oder einer verwandten Art, die zuvor geprüft und für hinreichend homogen befunden wurden. (vergleiche Dokument TGP/10, Abschnitt 5.2 „Bestimmung des zulässigen Variationsniveaus“).

### **9.2 Zusammenfassung**

9.2.1 Dokument TGP/10 erläutert, daß das Verfahren der Standardabweichung (SD) angewandt werden kann, wenn das Abweicherverfahren für die Prüfung der Homogenität nicht geeignet ist. Ferner sagt es folgendes in bezug auf die Bestimmung des zulässigen Variationsniveaus aus:

#### „5.2 Bestimmung des zulässigen Variationsniveaus

5.2.1.1 Der Vergleich zwischen einer Kandidatensorte und vergleichbaren Sorten wird aufgrund der Standardabweichungen vorgenommen, die aufgrund der Erfassungen an Einzelpflanzen errechnet werden. Die UPOV hat mehrere statistische Verfahren für die Behandlung der Homogenität bei gemessenen quantitativen Merkmalen vorgeschlagen. Der Vergleich zwischen einer Kandidatensorte und vergleichbaren Sorten wird aufgrund der mittels Erfassungen an Einzelpflanzen errechneten Standardabweichungen vorgenommen. Eine Methode, die die Variationen zwischen den Jahren berücksichtigt, ist die Methode des Kombinierten Homogenitätskriteriums über mehrere Jahre (COYU). Der Vergleich zwischen einer Kandidatensorte und vergleichbaren Sorten wird aufgrund der mittels Erfassungen an Einzelpflanzen errechneten Standardabweichungen vorgenommen. Dieses COYU-Verfahren berechnet eine Toleranzgrenze aufgrund bereits bekannter vergleichbarer Sorten, d. h. die Homogenität wird anhand einer relativen Toleranzgrenze geprüft, die auf Sorten im gleichen Anbauversuch mit vergleichbarer Merkmalsausprägung beruht.“

9.2.2 Die Homogenität ist häufig mit der Ausprägung eines Merkmals verbunden. So besteht z. B. bei einigen Arten bei Sorten mit größeren Pflanzen die Tendenz einer geringeren Homogenität in der Größe als bei Sorten mit kleineren Pflanzen. Wird derselbe Standard auf alle Sorten angewandt, ist es möglich, daß einige davon sehr strikte Kriterien erfüllen müssen, während für andere Standards vorhanden sind, die leicht zu erfüllen sind. COYU geht dieses Problem an, indem es vor der Festlegung eines Standards die Beziehungen zwischen der

Homogenität – ermittelt durch zwischen den Pflanzen gemessene Standardabweichung – und der Ausprägung des Merkmals – gemessen anhand des Sortenmittelwerts – anpaßt.

9.2.3 Die Methode erstreckt sich u. a. auf das Einstufen von vergleichbaren **Error! Bookmark not defined.** und Kandidatensorten nach dem Mittelwert des Merkmals. Es wird die SD jeder Sorte genommen und der SD-Mittelwert der ähnlichsten Sorten subtrahiert. Dieses Verfahren ermöglicht für jede Sorte eine Messung ihrer Homogenität, ausgedrückt in bezug auf diejenige vergleichbarer Sorten. Der Begriff „vergleichbare **Error! Bookmark not defined.** Sorte“ bezieht sich hier auf etablierte Sorten, die in den Anbauversuch einbezogen wurden und eine vergleichbare Ausprägung der in Prüfung befindlichen Merkmale haben.

9.2.4 Die Ergebnisse für jedes Jahr werden kombiniert, indem eine Sorte-x-Jahre-Tabelle der bereinigten SD erstellt und eine Varianzanalyse angewandt wird. Der bereinigte SD-Mittelwert für die Kandidatensorte wird mit dem Mittelwert für die vergleichbare **Error! Bookmark not defined.** Sorte verglichen, indem ein Standard-t-Test angewandt wird.

9.2.5 COYU vergleicht in der Tat die Homogenität einer Kandidatensorte mit derjenigen der in bezug auf das zu bewertende Merkmal ähnlichsten vergleichbaren **Error! Bookmark not defined.** Sorte. Die Vorteile von COYU liegen vor allem darin, daß alle Sorten auf der gleichen Grundlage verglichen und die Informationen von mehreren Prüfungsjahren in ein einziges Kriterium kombiniert werden können.

### **9.3 Einleitung**

9.3.1 Die Homogenität wird manchmal durch die Messung von einzelnen Merkmalen und Berechnung der Standardabweichung der Messungen an Einzelpflanzen innerhalb einer Parzelle berechnet. Die durchschnittlichen SD für alle Wiederholungen werden berechnet, um für jede Sorte in einem Anbauversuch einen einzigen Homogenitätsmaßstab zu ergeben.

9.3.2 Dieses Dokument umreißt ein Verfahren, das als das kombinierte Homogenitätskriterium über mehrere Jahre (COYU) bekannt ist. COYU prüft die Homogenität einer Sorte in bezug auf vergleichbare **Error! Bookmark not defined.** Sorten aufgrund der SD aus Anbauversuchen über mehrere Jahre. Eine Eigenschaft dieses Verfahrens ist, daß es die möglichen Beziehungen zwischen der Ausprägung eines Merkmals und der Homogenität berücksichtigt.

9.3.3 Dieser Abschnitte beschreibt

- die dem Verfahren COYU zugrundeliegenden Grundsätze;
- die UPOV-Empfehlungen zur Anwendung von COYU auf einzelne Arten;
- mathematische Einzelheiten über das Verfahren mit einem Beispiel für seine Anwendung;
- die für die Anwendung des Verfahrens verfügbare Computersoftware.

## 9.4 Das COYU-Kriterium

9.4.1 Das COYU-Verfahren besteht in einer Reihe von Schritten, die nachstehend aufgeführt sind. Diese werden auf jedes Merkmal der Reihe nach angewandt, Einzelheiten sind in Teil II Abschnitt 9.6 unten zu finden.

- Berechnung der SD innerhalb der Parzellen für jede Sorte in jedem Jahr.
- Umwandlung der SD durch Addieren von 1 und Konvertieren in natürliche Logarithmen.
- Schätzung der Beziehung zwischen der SD und dem Mittelwert in jedem Jahr. Das Verfahren beruht auf den gleitenden Durchschnitts der log-SD der vergleichbaren **Error! Bookmark not defined.** Sorten, angeordnet nach deren Mittelwerten.
- Bereinigungen der log-SD der Kandidatensorte und der vergleichbaren **Error! Bookmark not defined.** Sorten aufgrund der geschätzten Beziehungen zwischen der SD und dem Mittelwert in jedem Jahr.
- Berechnung des Durchschnitts der bereinigten log-SD über die Jahre.
- Berechnung der höchstzulässigen SD (das Homogenitätskriterium). Diese benutzt eine Schätzung der Variabilität bei der Homogenität der vergleichbaren **Error! Bookmark not defined.** Sorten, die aus der Varianzanalyse der Sorte-x-Jahr-Tabelle der bereinigten log-SD abgeleitet wird.
- Vergleich der bereinigten log-SD der Kandidatensorten mit der höchstzulässigen SD.

9.4.2 Die Vorteile des COYU-Verfahrens sind:

- Es bietet ein Verfahren zur Bewertung der Homogenität, das von den zu prüfenden Sorten weitgehend unabhängig ist.
- Das Verfahren kombiniert Informationen aus mehreren Anbauversuchen, um ein einziges Homogenitätskriterium zu bilden.
- Die auf dem Verfahren beruhenden Entscheidungen sind voraussichtlich im Verlauf der Zeit stabil.
- Das statistische Modell, auf dem es beruht, spiegelt die hauptsächlichen Variationsquellen wider, die die Homogenität beeinflussen.
- Die Standards beruhen auf der Homogenität der vergleichbaren **Error! Bookmark not defined.** Sorten.

## 9.5 Anwendung von COYU

9.5.1 COYU wird für die Verwendung bei der Prüfung der Homogenität von Sorten empfohlen

- für quantitative Merkmale;
- wenn Erfassungen auf Einzelpflanzenbasis über zwei oder mehrere Jahre erfolgen;
- wenn es Unterschiede zwischen Pflanzen einer Sorte gibt, die vielmehr eine quantitative Variation als das Vorhandensein von Abweichern darstellen.

9.5.2 Eine Sorte gilt für ein Merkmal als homogen, wenn sein bereinigter log-SD-Mittelwert das Homogenitätskriterium nicht übersteigt.

9.5.3 Das Wahrscheinlichkeitsniveau „p“, das für die Bestimmung des Homogenitätskriteriums angewandt wird, hängt von der Pflanze ab. Die empfohlenen Wahrscheinlichkeitsniveaus sind in Abschnitt 9.11 angegeben.

9.5.4 Die Homogenitätsprüfung kann über zwei oder drei Jahre erfolgen. Wenn die Prüfung in der Regel über drei Jahre durchgeführt wird, ist es möglich, sich für eine frühzeitige Akzeptanz oder Zurückweisung einer Sorte anhand der geeigneten Auswahl der Wahrscheinlichkeitswerte zu entscheiden.

9.5.5 Es wird empfohlen, für die Schätzung der Varianz für die vergleichbaren **Error! Bookmark not defined.** Sorten in der COYU-Analyse mindestens 20 Freiheitsgrade zu haben. Dies entspricht 11 vergleichbaren **Error! Bookmark not defined.** Sorten für eine COYU-Prüfung über zwei Jahre und 8 vergleichbaren **Error! Bookmark not defined.** Sorten in Dreijahresprüfungen). In manchen Fällen gibt es möglicherweise nicht genügend vergleichbare **Error! Bookmark not defined.** Sorten, um die empfohlenen Mindestfreiheitsgrade zu erreichen. Für diese Fälle wird zur Zeit eine Anleitung ausgearbeitet.

## 9.6 Mathematische Einzelheiten

Schritt 1: Ableitung der Standardabweichung innerhalb der Parzelle

9.6.1 Die Standardabweichungen innerhalb der Parzelle für jede Sorte in jedem Jahr werden mittels des Durchschnitts der Standardabweichungen zwischen den Pflanzen in der Parzelle,  $SD_j$ , über die Wiederholungen berechnet:

$$SD_j = \sqrt{\frac{\sum_{i=1}^n (y_{ij} - \bar{y}_j)^2}{(n-1)}}$$

$$SD = \frac{\sum_{j=1}^r SD_j}{r}$$

wobei  $y_{ij}$  die Beobachtung an der i. Pflanze in der j. Parzelle,  $\bar{y}_j$  der Mittelwert der Beobachtungen aus der j. Parzelle,  $n_j$  die Anzahl der Pflanzen in jeder Parzelle und r die Anzahl Wiederholungen ist.

Schritt 2: Umwandlung der SD

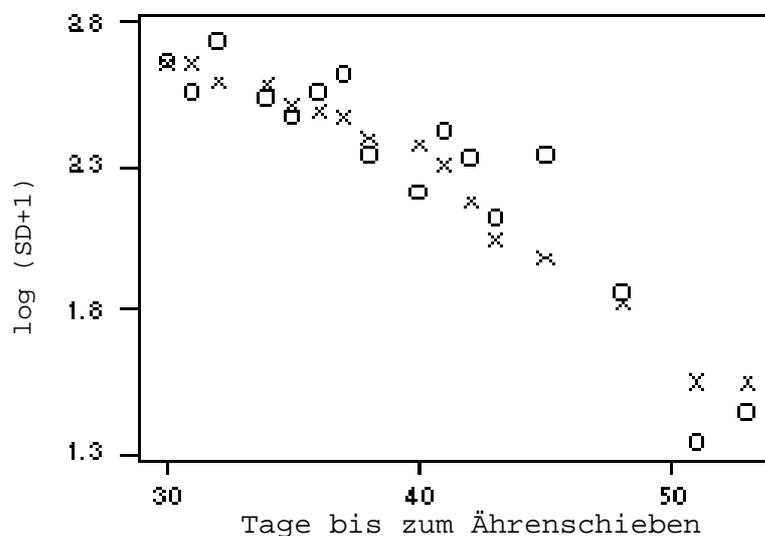
9.6.2 Die Umwandlung der SD durch Addieren von 1 und Konvertieren in natürliche Logarithmen. Mit dieser Umwandlung wird bezweckt, die SD für die statistische Analyse geeigneter zu machen.

Schritt 3: Schätzung der Beziehung zwischen der SD und dem Mittelwert in jedem Jahr

9.6.3 Für jedes Jahr wird die Form der durchschnittlichen Beziehung zwischen SD und Merkmalsmittelwert für die vergleichbaren **Error! Bookmark not defined.** Sorten separat geschätzt. Die Schätzungsmethode ist ein gleitender Durchschnitt von 9 Punkten. Die log-SD (die Y-Zufallsvariable) und die Mittelwerte (die X-Zufallsvariable) für jede Sorte werden zunächst gemäß den Werten des Mittelwerts eingeordnet. Für jeden Punkt ( $Y_i, X_i$ ) ist der Trendwert  $Y_i$  als Mittelwert der Werte  $Y_{i-4}, Y_{i-3}, \dots, Y_{i+4}$  zu nehmen, wobei  $i$  die Stufe des X-Wertes darstellt und  $Y_i$  der entsprechende Y-Wert ist. Für X-Werte mit den Stufen 1 und 2 wird für den Trendwert der Mittelwert der ersten drei Werte genommen. Im Falle des X-Wertes auf Stufe 3 wird der Mittelwert der ersten fünf Werte genommen, und für den X-Wert auf Stufe 4 wird der Mittelwert der ersten sieben Werte verwendet. Ein ähnliches Verfahren wird für die vier höchstrangigen X-Werte angewandt.

9.6.4 Dieses Verfahren wird für 16 Sorten anhand eines einfachen Beispiels in Abbildung 1 veranschaulicht. Die mit „0“ bezeichneten Punkte in der Abbildung 1a stellen für 16 Sorten die log-SD und die entsprechenden Sortenmittelwerte dar. Die mit „X“ bezeichneten Punkte sind die gleitenden Durchschnitte von 9 Punkten, die berechnet werden, indem für jede Sorte der Durchschnitt der log-SD der Sorte und der vier Sorten auf jeder Seite genommen wird. Der gleitende Durchschnitt an den äußeren Enden beruht auf dem Mittelwert von 3, 5 oder 7 Werten.

**Abbildung 1: Assoziation zwischen SD und Mittelwert – Tage bis zum Ährenschieben bei Sorten von Knaulgras** (das Symbol O bedeutet erfaßte SD, das Symbol X den gleitenden SD-Durchschnitt)



Schritt 4: Bereinigung der umgewandelten SD-Werte aufgrund der geschätzten Beziehung zwischen SD und Mittelwert

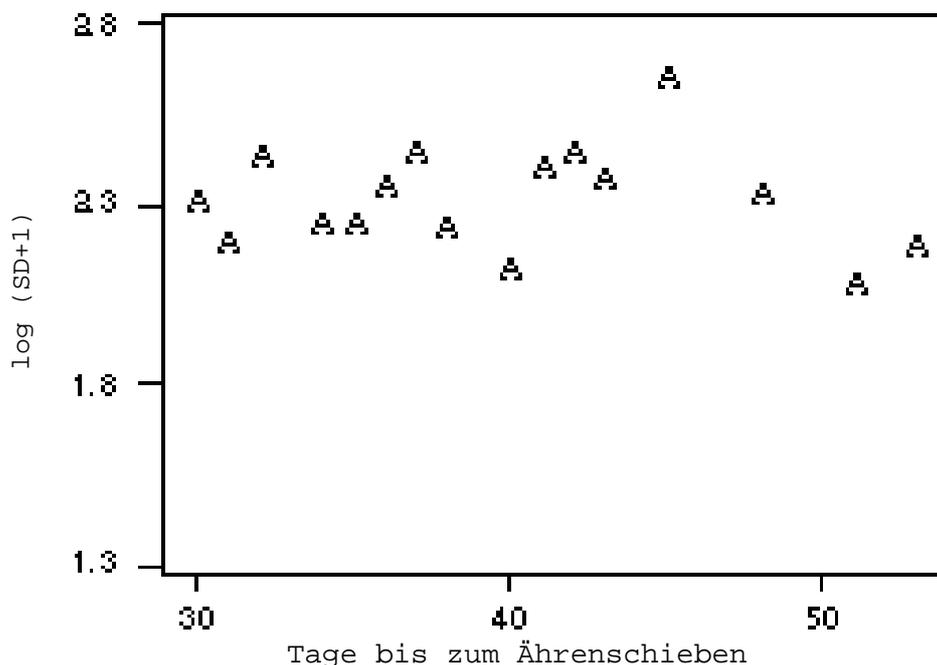
9.6.5 Sind die Trendwerte für die vergleichbaren **Error! Bookmark not defined.** Sorten festgelegt, dann werden die Trendwerte für die Kandidatensorten geschätzt, und zwar durch Anwendung linearer Interpolation zwischen den Trendwerten der nächstgelegenen zwei vergleichbaren **Error! Bookmark not defined.** Sorten, wie durch die Mittelwerte ihrer Merkmale definiert. Wenn also die Trendwerte für die beiden vergleichbaren **Error! Bookmark not defined.** Sorten auf jeder Seite der Kandidatensorte  $T_i$  und  $T_{i+1}$  sind, und wenn der beobachtete Wert für die Kandidatensorte  $X_c$  ist, wobei  $X_i \leq X_c \leq X_{i+1}$  ist, dann wird der Trendwert  $T_c$  für die Kandidatensorte angegeben durch

$$T_c = \frac{(X_c - X_i)T_{i+1} + (X_{i+1} - X_c)T_i}{X_{i+1} - X_i}$$

9.6.6 Um die SD in bezug auf ihre Beziehung zu dem Merkmalsmittelwert zu bereinigen, werden die geschätzten Trendwerte von den umgewandelten SD subtrahiert und der Gesamtmittelwert wieder zurückaddiert.

9.6.7 Die Ergebnisse für das einfache Beispiel mit 16 Sorten sind in Abbildung 2 dargestellt.

**Abbildung 2: Bereinigung bezüglich der Assoziation zwischen SD und Mittelwert – Tage bis zum Ährenschieben bei Sorten von Knaulgras (Symbol A bedeutet die bereinigte SD)**



#### Schritt 5: Berechnung des Homogenitätskriteriums

9.6.8 Die Variabilität in der Homogenität der vergleichbaren **Error! Bookmark not defined.** Sorten wird geschätzt, indem eine Einweg-Varianzanalyse für die log-SD angewandt wird, d. h. mit den Jahren als klassifizierendem Faktor. Die Variabilität (V) wird anhand des Restwertes in dieser Varianzanalyse geschätzt.

9.6.9 Die höchstzulässige Standardabweichung (das Homogenitätskriterium) auf der Grundlage von Anbauversuchen über k Jahre ist wie folgt:

$$UC_p = SD_r + t_p \sqrt{V \left( \frac{1}{k} + \frac{1}{Rk} \right)}$$

wobei

$SD_r$  der Mittelwert der bereinigten log-SD für die vergleichbaren **Error! Bookmark not defined.** Sorten, V die Varianz der bereinigten log-SD nach Beseitigung der Jahreseffekte,  $t_p$  der einseitige t-Wert für die Wahrscheinlichkeit p mit Freiheitsgraden wie für V, k die Anzahl Jahre und R die Anzahl vergleichbarer **Error! Bookmark not defined.** Sorten ist.

### 9.7 Frühzeitige Entscheidungen für eine Dreijahresprüfung

9.7.1 Die Entscheidungen über die Homogenität können je nach Pflanze nach zwei oder drei Jahren getroffen werden. Wenn COYU in der Regel über drei Jahre durchgeführt wird, ist es möglich, eine Sorte anhand einer geeigneten Auswahl der Wahrscheinlichkeitswerte frühzeitig zu akzeptieren oder zurückzuweisen.

9.7.2 Das Wahrscheinlichkeitsniveau für die frühzeitige Zurückweisung einer Kandidatensorte nach zwei Jahren sollte gleich sein wie für die volle Dreijahresprüfung. Wenn beispielsweise die Dreijahres-COYU-Prüfung anhand eines Wahrscheinlichkeitsniveaus von 0,2 % durchgeführt wird, kann eine Kandidatensorte nach zwei Jahren zurückgewiesen werden, wenn ihre Homogenität das COYU-Kriterium mit dem Wahrscheinlichkeitsniveau 0,2 % übersteigt.

9.7.3 Das Wahrscheinlichkeitsniveau für die frühe Akzeptanz einer Kandidatensorte nach zwei Jahren sollte höher sein als für die volle Dreijahresprüfung. Wenn beispielsweise die Dreijahres-COYU-Prüfung anhand eines Wahrscheinlichkeitsniveaus von 0,2 % durchgeführt wird, kann eine Kandidatensorte nach zwei Jahren akzeptiert werden, wenn ihre Homogenität das COYU-Kriterium mit dem Wahrscheinlichkeitsniveau 2 % nicht übersteigt.

9.7.4 Einige Sorten werden nach zwei Jahren möglicherweise nicht zurückgewiesen oder akzeptiert. In dem in den Absätzen 26 und 27 oben dargelegten Beispiel könnte eine Sorte eine Homogenität haben, die das COYU-Kriterium mit dem Wahrscheinlichkeitsniveau 2 %, jedoch nicht das Kriterium mit dem Wahrscheinlichkeitsniveau 0,2 % übersteigt. In diesem Fall sollten diese Sorten nach drei Jahren erneut geprüft werden.

## 9.8 Beispiel für COYU-Berechnungen

9.8.1 Ein Beispiel für die Anwendung von COYU wird hier gegeben, um die damit verbundenen Berechnungen zu verdeutlichen. Das Beispiel betrifft die Punktzahlen für die Tage bis zum Ährenschieben für Deutsches Weidelgras über drei Jahre für 11 vergleichbare **Error! Bookmark not defined.** Sorten (R1 bis R11) und eine Kandidatensorte (C1). Die Daten sind in Tabelle 1 aufgeführt.

**Tabelle 1: Beispiel eines Datensatzes – Tage bis zum Ährenschieben bei Deutschem Weidelgras**

Sorte	Merkmalsmittelwerte			SD innerhalb der Parzellen			log(SD+1)		
	Jahr 1	Jahr 2	Jahr 3	Jahr 1	Jahr 2	Jahr 3	Jahr 1	Jahr 2	Jahr 3
R1	38	41	35	8,5	8,8	9,4	2,25	2,28	2,34
R2	63	68	61	8,1	7,6	6,7	2,21	2,15	2,04
R3	69	71	64	9,9	7,6	5,9	2,39	2,15	1,93
R4	71	75	67	10,2	6,6	6,5	2,42	2,03	2,01
R5	69	78	69	11,2	7,5	5,9	2,50	2,14	1,93
R6	74	77	71	9,8	5,4	7,4	2,38	1,86	2,13
R7	76	79	70	10,7	7,6	4,8	2,46	2,15	1,76
R8	75	80	73	10,9	4,1	5,7	2,48	1,63	1,90
R9	78	81	75	11,6	7,4	9,1	2,53	2,13	2,31
R10	79	80	75	9,4	7,6	8,5	2,34	2,15	2,25
R11	76	85	79	9,2	4,8	7,4	2,32	1,76	2,13
C1	52	56	48	8,2	8,4	8,1	2,22	2,24	2,21

9.8.2 Die Berechnungen für die Bereinigung der SD im Jahr 1 sind in Tabelle 2 angegeben. Den Trendwert der Kandidatensorte erhält man durch Interpolation zwischen den Werten für die Sorten R1 und R2, da der Merkmalsmittelwert für C1 (d. h. 52) zwischen den Mittelwerten für R1 und R2 (d. h. 38 und 63) liegt. Das heißt

$$T_c = \frac{(X_c - X_i)T_{i+1} + (X_{i+1} - X_c)T_i}{X_{i+1} - X_i} = \frac{(52 - 38) \times 2,28 + (63 - 52) \times 2,28}{63 - 38} = 2,28$$

**Tabelle 2: Beispiel für einen Datensatz – Berechnung des bereinigten  $\log(SD+1)$  für das Jahr 1**

Sorte	geordneter Mittelwert (X)	$\log(SD+1)$ (Y)	Trendwert T	bereinigte $\log(SD+1)$
R1	38	2,25	$(2,25 + 2,21 + 2,39)/3 = 2,28$	$2,25 - 2,28 + 2,39 = 2,36$
R2	63	2,21	$(2,25 + 2,21 + 2,39)/3 = 2,28$	$2,21 - 2,28 + 2,39 = 2,32$
R3	69	2,39	$(2,25 + \dots + 2,42)/5 = 2,35$	$2,39 - 2,35 + 2,39 = 2,42$
R5	69	2,50	$(2,25 + \dots + 2,48)/7 = 2,38$	$2,50 - 2,38 + 2,39 = 2,52$
R4	71	2,42	$(2,25 + \dots + 2,32)/9 = 2,38$	$2,42 - 2,38 + 2,39 = 2,43$
R6	74	2,38	$(2,21 + \dots + 2,53)/9 = 2,41$	$2,38 - 2,41 + 2,39 = 2,36$
R8	75	2,48	$(2,39 + \dots + 2,34)/9 = 2,42$	$2,48 - 2,42 + 2,39 = 2,44$
R7	76	2,46	$(2,42 + \dots + 2,34)/7 = 2,42$	$2,46 - 2,42 + 2,39 = 2,43$
R11	76	2,32	$(2,48 + \dots + 2,34)/5 = 2,43$	$2,32 - 2,43 + 2,39 = 2,28$
R9	78	2,53	$(2,32 + 2,53 + 2,34)/3 = 2,40$	$2,53 - 2,40 + 2,39 = 2,52$
R10	79	2,34	$(2,32 + 2,53 + 2,34)/3 = 2,40$	$2,34 - 2,40 + 2,39 = 2,33$
Mittelwert	70	2,39		
C1	52	2,22	2,28	$2,22 - 2,28 + 2,39 = 2,32$

9.8.3 Die Ergebnisse der Bereinigung für alle drei Jahre sind in Tabelle 3 enthalten.

**Tabelle 3: Beispiel eines Datensatzes – bereinigte  $\log(SD+1)$  für alle drei Jahre mit den Mittelwerten über die Jahre**

Sorte	Mittelwert über die Jahre		bereinigte $\log(SD+1)$		
	Merkmalsmittelwert	bereinigte $\log(SD+1)$	Jahr 1	Jahr 2	Jahr 3
R1	38	2,26	2,36	2,13	2,30
R2	64	2,10	2,32	2,00	2,00
R3	68	2,16	2,42	2,10	1,95
R4	71	2,15	2,43	1,96	2,06
R5	72	2,20	2,52	2,14	1,96
R6	74	2,12	2,36	1,84	2,16
R7	75	2,14	2,43	2,19	1,80
R8	76	2,02	2,44	1,70	1,91
R9	78	2,30	2,52	2,16	2,24
R10	78	2,22	2,33	2,23	2,09
R11	80	2,01	2,28	1,78	1,96
Mittelwert	70	2,15	2,40	2,02	2,04
C1	52	2,19	2,32	2,08	2,17

9.8.4 Die Varianzanalysetabelle für die bereinigten  $\log$ -SD ist in Tabelle 4 angegeben (nur aufgrund der vergleichbaren **Error! Bookmark not defined.** Sorten). Die Variabilität bei der Homogenität der vergleichbaren **Error! Bookmark not defined.** Sorten wird auf dieser Grundlage geschätzt ( $V=0,0202$ ).

**Tabelle 4: Beispiel eines Datensatzes – Varianzanalysetabelle für die bereinigte  $\log(\text{SD}+1)$**

Ursache	Freiheits- grade	Summen der Quadrate	Mittelwert- quadrate
Jahr	2	1,0196	0,5098
Sorten innerhalb der Jahre (=Residuum)	30	0,6060	<b>0,0202</b>
Insgesamt	32	1,6256	

9.8.5 Das Homogenitätskriterium für ein Wahrscheinlichkeitsniveau von 0,2 % wird wie folgt berechnet:

$$UC_p = SD_r + t_p \sqrt{V \left( \frac{1}{k} + \frac{1}{Rk} \right)} = 2.15 + 3.118 \times \sqrt{0.0202 \times \left( \frac{1}{3} + \frac{1}{3 \times 11} \right)} = 2.42$$

wobei  $t_p$  der Student-t-Tabelle entnommen ist mit  $p = 0,002$  (einseitig) und 30 Freiheitsgraden.

9.8.6 Sorten mit mittelwertbereinigter  $\log(\text{SD} + 1)$  geringer als oder gleich 2,42 können für dieses Merkmal als homogen betrachtet werden. Die Kandidatensorte C1 erfüllt dieses Kriterium.

## **9.9 Umsetzung von COYU**

Das COYU-Kriterium kann unter Inanspruchnahme des COYU-Moduls des DUST-Softwarepakets für statistische Analysen von DUS-Daten verwendet werden, das erhältlich ist von Frau Sally Watson, (Email: [info@afbini.gov.uk](mailto:info@afbini.gov.uk)) oder unter <http://www.afbini.gov.uk/dustnt.htm>.

## **9.10 COYU-Software**

### 9.10.1 DUST-Computerprogramm

9.10.1.1 Die Hauptdatenausgabe des DUST-COYU-Programms ist in Tabelle A1 dargestellt, in der die Ergebnisse der Analysen der SD innerhalb der Parzellen für 49 Sorten von Deutschem Weidelgras zusammengefaßt sind, die über eine Periode von drei Jahren bewertet wurden. Weitere Ergebnisse sind in Tabelle A2 enthalten, in der Einzelheiten über die Analyse eines einzelnen Merkmals, den Zeitpunkt des Ährenschiebens, angegeben sind. Es ist zu beachten, daß die angegebene Varianzanalysetabelle eine weitere Varianzursache hat; die Variationsursache, V, der bereinigten  $\log$ -SD wird durch Kombinieren der Variation für die Sorte und die Restursachen berechnet.

9.10.1.2 In Tabelle A1 ist die bereinigte SD für jede Sorte als Prozentsatz des SD-Mittelwerts für alle vergleichbaren **Error! Bookmark not defined.** Sorten ausgedrückt. Mit der Zahl 100 wird eine Sorte mit durchschnittlicher Homogenität angegeben; eine Sorte

mit einem geringeren Wert als 100 zeigt gute Homogenität an; eine Sorte mit einem viel größeren Wert als 100 deutet schlechte Homogenität bei diesem Merkmal an. Mangelnde Homogenität in einem Merkmal wird häufig durch den Beweis mangelnder Homogenität in verwandten Merkmalen bestätigt.

9.10.1.3 Die Zeichen „\*“ und „+“ rechts von den Prozentzahlen identifizieren Sorten, deren SD das COYU-Kriterium nach 3 bzw. 2 Jahren überschreiten. Das Zeichen „:“ gibt an, daß die Homogenität nach zwei Jahren noch nicht akzeptierbar ist und daß die Prüfung der Sorte für ein weiteres Jahr in Betracht gezogen werden sollte. Es ist zu beachten, daß für dieses Beispiel ein Wahrscheinlichkeitsniveau von 0,2 % für die Dreijahresprüfung angewandt wird. Für frühzeitige Entscheidungen nach zwei Jahren werden die Wahrscheinlichkeitsniveaus 2 % und 0,2 % für die Akzeptanz bzw. Zurückweisung der Sorten angewandt. Alle Kandidatensorten wiesen aufgrund des COYU-Kriteriums für die acht Merkmale eine zulässige Homogenität aus.

9.10.1.4 Die Zahlen rechts von den Prozentsätzen geben die Anzahl Jahre an, in denen das Kriterium der Homogenität innerhalb eines Jahres übertroffen wurde. Dieses Kriterium wurde nun durch COYU ersetzt.

9.10.1.5 Das Programm funktioniert mit einem vollständigen Datensatz oder akzeptiert einige fehlende Werte, z. B. wenn eine Sorte in einem Jahre nicht vorhanden ist.

**Tabelle A1: Beispiel einer zusammenfassenden Datenausgabe des COYU-Programms**

\*\*\*\* ZUSAMMENFASSUNG DER HOMOGENITÄTSANALYSE  
 ÜBER DIE JAHRE \*\*\*\*

STANDARDABWEICHUNGEN INNERHALB DER PARZELLE ALS

PROZENTMITTELWERT DER SD DER VERGLEICHSSORTEN

	5	60	8	10	11	
R1	100	100	95 1	100	97	97
R2	105	106	98	99	104	101
R3	97	103	92 1	103	96	98
R4	102	99	118 2	105	101	101
R5	102	99	116 3	95	104	110
R6	103	102	101	99	97	104
R7	100	95	118 2	102 1	98	99
R8	97	98	84	95	97	93
R9	97	105	87	99	101	99
R10	104	100	96	105 1	96	102
R11	99	96	112	99	101	98
R12	100	97	99 1	103	105	106
R13	95	96	101	100	96	101
R14	105	103	90	97	101	97
R15	102	100 1	89	105	105 1	101
R16	99	98	92 1	98	102	98
R17	97	101	98	101	101	95
R18	99	97	96	96	102	99
R19	103	101	105	102	100	98
R20	104	99	93	91	100	102
R21	97	94	103	97	100	102
R22	101	110*1	112	107 1	103 1	101
R23	94	101	107	99	104	97
R24	99	97	95	99	100	103
R25	104 1	103	93 1	99	101	96
R26	98	97	111 2	96	102 1	106
R27	102	99	106 1	99	103	107
R28	101	106	90	95	101	101
R29	101	105	83	102	94	93
R30	99	96	97	99	95	100
R31	99	102	107	107 1	102	99
R32	98	93	111 2	102	98	103
R33	104	102 1	107 1	103	100	97
R34	95	94	82	95	97	96
R35	100	102	95	100	99	94
R36	99	98	111 1	99	100	103
R37	100	107 1	107	101	100	107
R38	95	97	102	107 1	97	101
R39	99	99	90	98	101	100
R40	104	102	112 1	100	101	97
C1	100 1	106	113 2	104 1	106 1	106
C2	103	101	98	97	101	109
C3	97	93	118 2	98	99	109
C4	102	101	106	103	99	101
C5	100	104	99	103	100	107
C6	101	102	103	100	103	107
C7	96	98	106	97	102	103
C8	101	105 1	116 2	103	103	93
C9	99	99	90 2	91	97	98

MERKMALSSCHLÜSSEL:

5	FRÜHJAHRSHÖHE	60	NATÜRLICHE
8	DATUM ÄHRENSCHIEBEN	10	HÖHE BEI
11	BREITE BEI		
15	BREITE DES	14	LÄNGE DES
		24	ÄHRENLÄNGE

SYMBOLE

\* - SD HÖHER ALS ÜBERJAHRESKRITERIUM NACH 3 J.  
 -----SD HÖHER ALS ÜBERJAHRESKRITERIUM NACH 2 J.  
 -----SD HÖHER ALS ÜBERJAHRESKRITERIUM NACH 1 J.

1, 2, 3 - ANZAHL FÄLLE IN DENEN SD INNERHALB DER JA  
 HRE ÜBER DAS ÜBERJAHRESKRITERIUM ÜBERSCREITET

TGP/8/1 Draft 16: TEIL II: 9: KOMBINIERTES HOMOGENITÄTSKRITERIUM ÜBER MEHRERE JAHRE (COYU)

Seite 126

**Tabelle A2: Beispiel einer weiteren Datenausgabe der DUST-Software für das Merkmal „Zeitpunkt des Ährenschiebens“ (Merkmal 8)**

\*\*\*\* HOMOGENITÄTSANALYSE VON STANDARDABWEICHUNGEN ZWISCHEN PFLANZEN (SD) :

SORTE	ÜBER JAHRE			INDIVIDUELLE JAHRE								
	MERKMALS- MITTELWERT	BER. LOG SD	NICHT BER. LOG SD	MERKMALSMITTELWERT-			LOG(SD+1)			BER. LOG(S)		
				88	89	90	88	89	90	88	89	9
VERGLEICHSSO												
R5	38.47	1.823	2.179	39.07	41.21	35.12	2.02	2.18	2.34X	1.73	1.78	1.96
R5	50.14	2.315	2.671	48.19	53.69	48.54	2.52X	2.74X	2.76X	2.23	2.33	2.39
R16	59.03	1.833	2.179	57.25	63.33	56.50	2.28X	2.24	2.01	1.96	1.73	1.81
R26	63.44	2.206	2.460	61.00	66.53	62.81	2.50X	2.75X	2.13	2.18	2.33	2.11
R9	63.99	1.739	1.994	62.92	68.32	60.72	2.21	2.03	1.74	1.96	1.64	1.62
R12	66.12	1.964	2.086	67.89	65.35	65.12	2.07	2.58X	1.60	1.97	2.14	1.78
R33	67.58	2.124	2.254	66.66	71.54	64.53	2.55X	2.26	1.95	2.32	1.92	2.12
R1	67.87	1.880	1.989	69.07	70.64	63.90	1.60	2.45X	1.93	1.60	2.08	1.96
R20	68.74	1.853	1.893	67.17	74.31	64.74	2.05	1.95	1.68	1.92	1.75	1.89
R25	68.82	1.853	1.905	68.28	72.38	65.81	1.83	2.39X	1.49	1.75	2.09	1.72
R18	69.80	1.899	1.853	68.61	75.22	65.58	1.88	1.84	1.84	1.82	1.80	2.08
R30	70.53	1.919	1.864	70.36	75.08	66.15	2.04	1.84	1.71	2.00	1.78	1.98
R13	70.63	2.005	2.000	70.23	75.00	66.66	1.97	2.03	2.01	1.91	1.86	2.24
R32	71.49	2.197	2.238	70.03	74.98	69.44	2.32X	2.45X	1.94	2.31	2.27	2.01
R34	72.09	1.630	1.545	71.32	77.35	67.59	1.57	1.49	1.58	1.54	1.58	1.78
R40	72.24	2.222	2.178	72.71	75.07	68.95	2.25X	2.26	2.03	2.29	2.16	2.22
R23	72.40	2.122	2.058	69.72	78.39	69.10	2.11	2.14	1.93	2.16	2.14	2.06
R29	72.66	1.657	1.580	73.13	75.80	69.04	1.46	1.63	1.65	1.47	1.69	1.81
R7	73.19	2.341	2.342	72.23	75.80	71.52	2.62X	2.30X	2.10	2.61	2.30	2.11
R24	73.19	1.888	1.796	74.00	76.37	69.20	1.62	1.84	1.93	1.71	1.91	2.04
R19	73.65	2.083	2.049	73.32	76.06	71.57	1.96	2.05	2.14	1.96	2.13	2.16
R2	73.85	1.946	1.897	72.98	78.16	70.42	1.76	1.96	1.97	1.79	2.02	2.03
R31	74.23	2.119	2.012	73.73	78.23	70.71	2.05	1.86	2.13	2.25	1.94	2.17
R37	74.38	2.132	2.020	74.87	76.95	71.32	1.97	2.04	2.04	2.23	2.11	2.06
R11	74.60	2.224	2.150	73.87	78.07	71.87	2.21	2.08	2.16	2.36	2.10	2.21
R38	74.76	2.029	1.916	76.11	78.24	69.93	1.84	2.15	1.75	1.98	2.24	1.87
R8	74.83	1.677	1.593	74.27	78.77	71.45	1.62	1.55	1.61	1.75	1.64	1.64
R15	75.54	1.760	1.682	75.72	78.68	72.22	1.53	1.79	1.73	1.64	1.84	1.80
R10	75.64	1.915	1.847	73.47	79.24	74.23	1.87	1.66	2.00	1.99	1.78	1.98
R22	75.68	2.228	2.133	74.57	79.17	73.32	2.18	2.21	2.01	2.40	2.26	2.03
R14	75.84	1.797	1.688	74.53	79.56	73.43	1.54	1.63	1.90	1.70	1.76	1.93
R17	76.13	1.942	1.832	75.34	79.09	73.96	1.65	2.04	1.81	1.90	2.10	1.83
R39	76.83	1.781	1.676	75.49	80.50	74.50	1.56	1.51	1.96	1.72	1.70	1.92
R35	77.22	1.886	1.773	76.67	80.85	74.15	1.73	1.67	1.92	1.88	1.85	1.93
R4	77.78	2.349	2.268	76.80	81.22	75.33	2.36X	2.13	2.31X	2.52	2.33	2.20
R36	77.98	2.209	2.173	78.97	79.85	75.11	2.13	2.15	2.25X	2.24	2.21	2.18
R6	78.73	2.009	1.935	77.53	82.88	75.78	2.00	1.75	2.06	2.03	2.09	1.91
R27	78.78	2.116	2.098	77.61	80.03	78.69	1.80	2.25	2.24X	1.87	2.39	2.09
R28	79.41	1.785	1.722	78.28	81.99	77.97	1.68	1.43	2.05	1.79	1.67	1.89
R21	80.52	2.045	1.950	77.43	85.02	79.11	1.98	1.75	2.13	2.07	2.09	1.98
KANDIDATENSORTE												
C1	64.03	2.252	2.438	63.85	63.33	64.92	2.49X	2.81X	2.02	2.25	2.29	2.21
C2	86.11	1.940	1.837	84.83	88.63	84.85	1.79	1.71	2.01	1.90	2.05	1.87
C3	82.04	2.349	2.248	82.26	87.45	76.40	2.37X	2.03	2.35X	2.48	2.37	2.20
C4	78.63	2.104	2.033	78.01	82.17	75.72	2.05	2.01	2.04	2.15	2.27	1.90
C5	72.99	1.973	1.869	71.98	79.40	67.59	1.95	1.78	1.88	1.93	1.90	2.08
C6	83.29	2.050	1.947	84.10	85.57	80.21	2.05	1.69	2.10	2.16	2.03	1.96
C7	83.90	2.100	1.997	84.12	87.99	79.60	1.93	1.95	2.11	2.04	2.29	1.97
C8	83.50	2.304	2.201	82.43	85.98	82.08	2.27X	2.00	2.34X	2.38	2.33	2.20
C9	51.89	1.788	2.157	52.35	55.77	47.56	1.83	2.34X	2.31X	1.52	1.91	1.93
MITTELWERT												
VERGL. SORTE	71.47	1.988		70.78	74.97	68.65	1.97	2.03	1.96	1.99	1.99	1.99
HOMOGENITÄTSKRITERIUM												
				WAHRSCHEINLICHKEITSNIVEAU								
3-JAHRESZURÜCKWEISUNG		2.383	0.002									
2-JAHRESZURÜCKWEISUNG		2.471	0.002									
2-JAHRESAKZEPTANZ		2.329	0.020									
**** VARIANZANALYSE DER BEREINIGTEN LOG(SD+1)												
	DF	MS	F	VERHÄLTNIS								
JAHRE	2	0.06239										
SORTEN	39	0.11440	5.1									
REST	78	0.02226										
GESAMT	119	0.05313										
SYMBOLS												
	* - SD HÖHER ALS ÜBERJAHRESHOMOGENITÄTSKRITERIUM NACH 3 JAHREN											
	+ - SD HÖHER ALS ÜBERJAHRESHOMOGENITÄTSKRITERIUM NACH 2 JAHREN											
	: - SD NOCH NICHT AKZEPTABEL GEMÄSS ÜBERJAHRESKRITERIUM NACH 2 JAHREN											
	X - SD HÖHER ALS 1,265FACHER MITTELWERT DER VERGLEICHSSORTEN											

### 9.11 Schemata für die Anwendung von COYU

Die folgenden vier Fälle stellen im allgemeinen die verschiedenen Situationen dar, die auftreten können, wenn COYU bei der DUS-Prüfung angewandt wird:

Schema A: Die Prüfung wird über zwei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach zwei Wachstumsperioden getroffen (eine Wachstumsperiode könnte ein Jahr sein und wird nachstehend als Periode bezeichnet)

Schema B: Die Prüfung wird über drei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen

Schema C: Die Prüfung wird über drei unabhängige Wachstumsperioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen, doch kann eine Sorte auch nach zwei Perioden akzeptiert werden

Schema D: Die Prüfung wird über drei unabhängige Perioden durchgeführt, und die Entscheidungen werden nach drei Perioden getroffen, doch kann eine Sorte auch nach zwei Perioden akzeptiert oder zurückgewiesen werden

Die Stadien, in denen die Entscheidungen in den Fällen A bis D getroffen werden, sind in den Abbildungen 1 bis 4 dargestellt. Diese verdeutlichen auch die verschiedenen Standard-Wahrscheinlichkeitsniveaus ( $p_{u2}$ ,  $p_{nu2}$  und  $p_{u3}$ ), die für die Berechnung des COYU-Kriteriums je nach Fall erforderlich sind. Diese werden wie folgt definiert:

<b>Wahrscheinlichkeitsniveau</b>	<b>verwendet, um zu entscheiden, ob eine Sorte</b>
$p_{u2}$	nach 2 Perioden homogen ist
$p_{md2}$	nach 2 Perioden nicht homogen ist
$p_{u3}$	nach 3 Perioden in einem Merkmal homogen ist

In den Abbildungen 1 bis 4 wird das COYU-Kriterium, das beispielsweise anhand des Wahrscheinlichkeitsniveaus  $p_{u2}$  berechnet wird, als  $UC_{p_{u2}}$  bezeichnet usw. Der Begriff „U“ stellt die mittelwertbereinigte  $\log(SD+1)$  einer Sorte für ein Merkmal dar.

Tabelle 1 faßt die verschiedenen Standard-Wahrscheinlichkeitsniveaus zusammen, die für die Berechnung des COYD- und des COYU-Kriteriums in jedem der Fälle A bis D erforderlich sind. In Fall B beispielsweise ist nur ein Wahrscheinlichkeitsniveau ( $p_{u3}$ ) erforderlich, während Fall C zwei davon benötigt ( $p_{u2}$ ,  $p_{u3}$ ).

FALL	COYU		
	$p_{u2}$	$p_{nu2}$	$p_{u3}$
A			
B			
C			
D			

Abbildung 1. COYU-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) im Fall A

COYU                      Entscheidung nach der 2. Periode

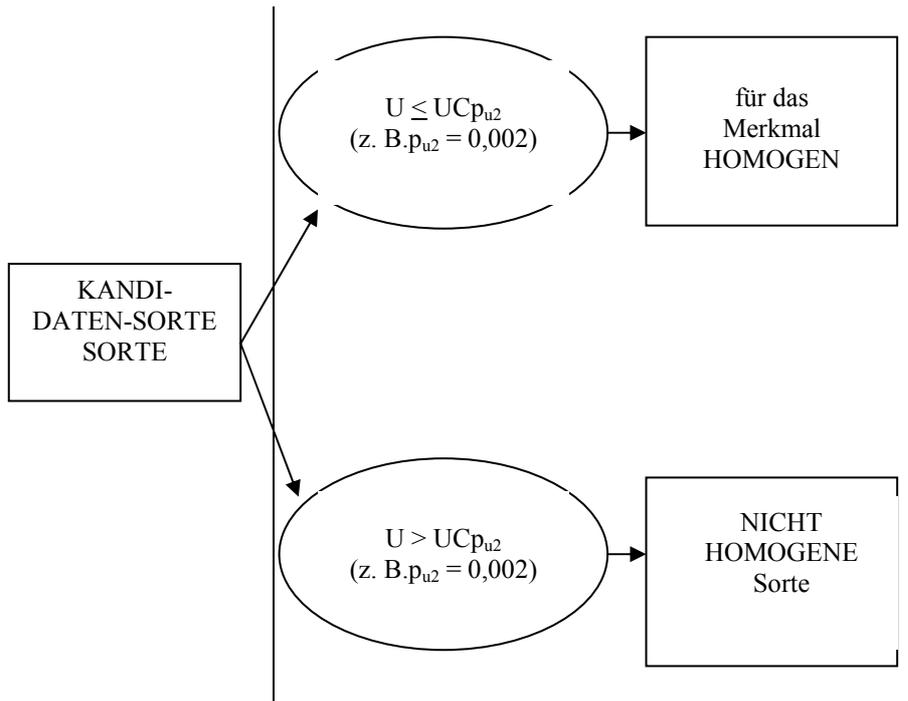
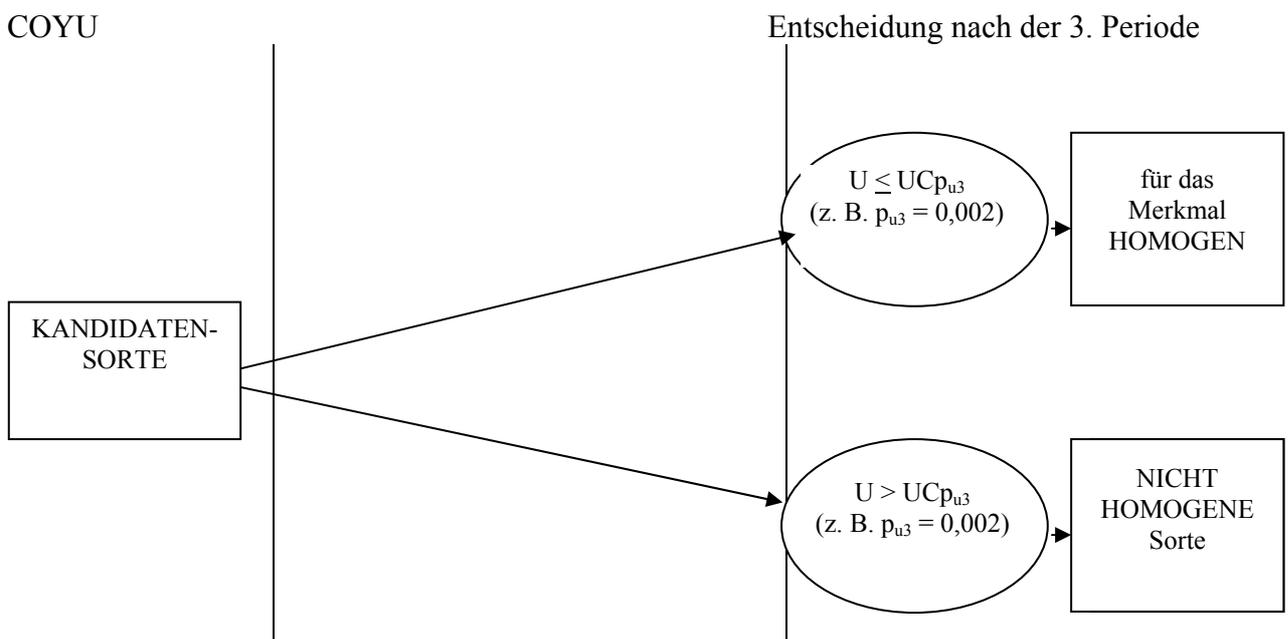


Abbildung 2. COYD- und COYU-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) im Fall B

COYU



ANMERKUNG:

„U“ ist die mittelwertbereinigte  $\log(SD+1)$  der Kandidatensorte für das Merkmal.

$UCp$  ist das auf dem Wahrscheinlichkeitsniveau  $p$  berechnete COYU-Kriterium.

Abbildung 3. COYU-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) im Fall C

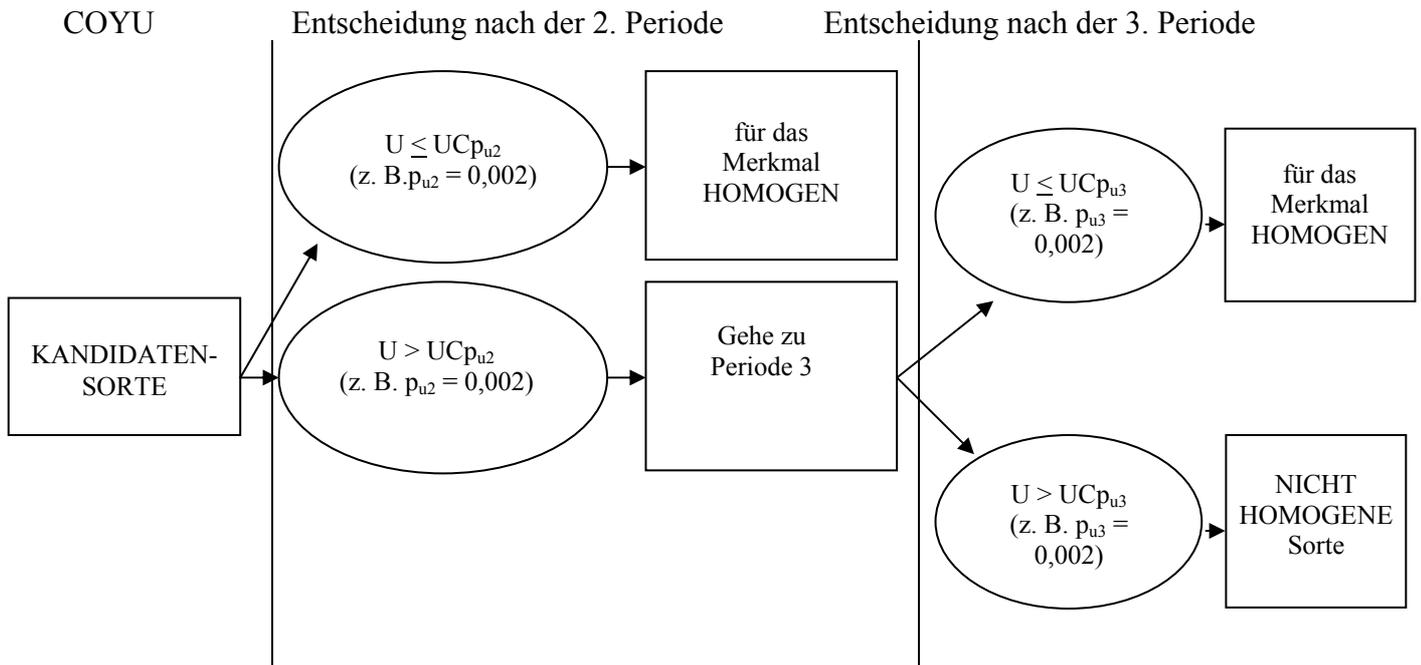
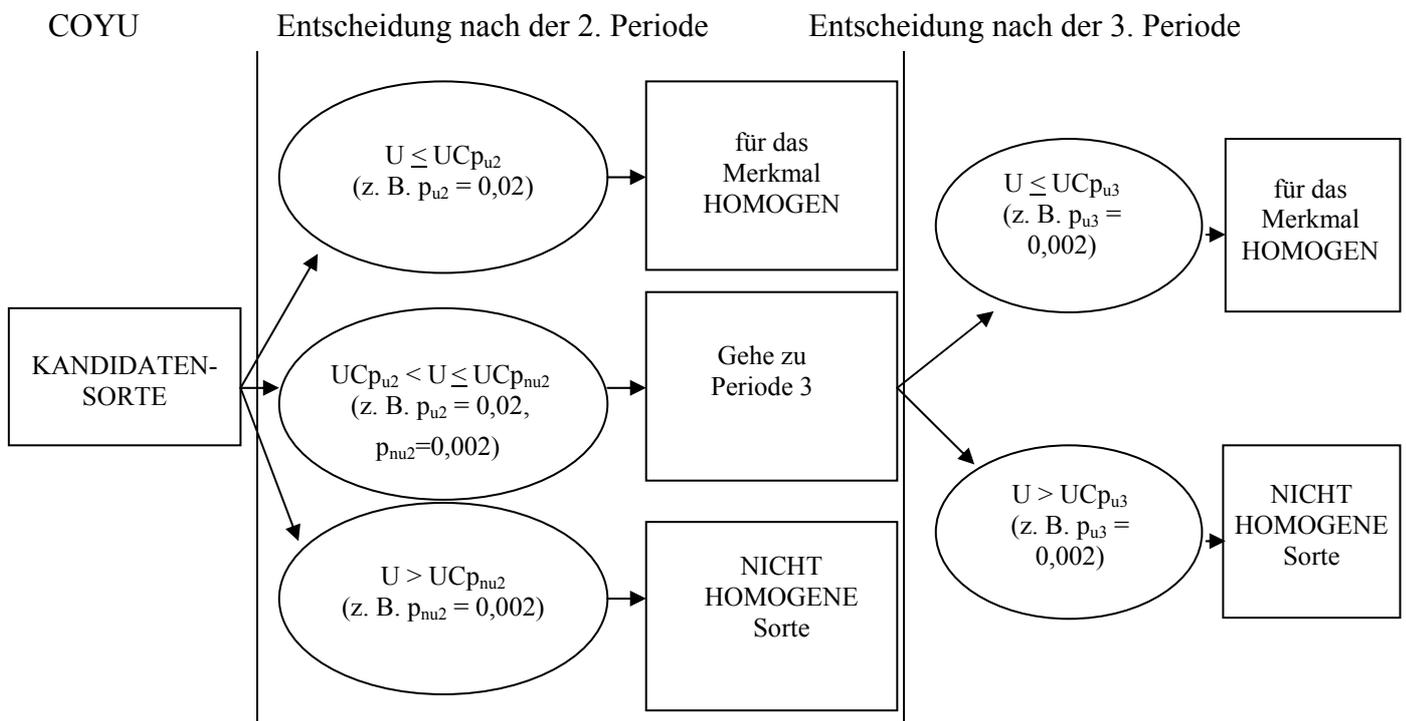


Abbildung 4. COYD- und COYU-Entscheidungen und Standard-Wahrscheinlichkeitsniveaus ( $p_i$ ) im Fall D



ANMERKUNG:

„U“ ist die mittelwertbereinigte  $\log(SD+1)$  der Kandidatensorte für das Merkmal.

UCp ist das auf dem Wahrscheinlichkeitsniveau  $p$  berechnete COYU-Kriterium.

## 10. PRÜFUNG DER HOMOGENITÄT ANHAND VON VERFAHREN DER RELATIVEN VARIANZ

### 10.1 Anwendung des Verfahrens der relativen Varianz

Die relative Varianz für ein bestimmtes Merkmal bezieht sich auf die Varianz der Kandidatensorte dividiert durch die durchschnittliche Varianz der vergleichbaren **Error! Bookmark not defined.** Sorten (i.e. Relative variance = variance of the candidate/average variance of the comparable **Error! Bookmark not defined.** varieties). Die Daten sollten normal verteilt sein. Die Verfahren der relative Varianz können auf jedes erfasste Merkmal angewandt werden, be idem es sich um eine kontinuierliche Variable handelt, unabhängig von der Vermehrungsweise der Sorte. Vergleichbare Sorten sind Sorten desselben Typs innerhalb derselben oder einer verwandten Art, die zuvor geprüft und für hinreichend homogen befunden wurden. (vergleiche Dokument TGP/10, Abschnitt 5.2 „Bestimmung des zulässigen Variationsniveaus“).

Bei fremdbefruchtenden Sorten lautet eine allgemeine Empfehlung in den UPOV-Prüfungsrichtlinien, 60 Messungen je Merkmal und Sorte vorzunehmen. Im wesentlichen ist das Varianzverhältnis gleich der Kenngröße F, und der tabellierte Wert von F bei  $P = 0,01$  unter  $df_1 = 60$  (Freiheitsgrade der Kandidatensorte) und  $df_2 = \infty$  (Freiheitsgrade der vergleichbaren **Error! Bookmark not defined.** Sorte(n)) ist 1,60.  $df_2 = \infty$  wird als vorsichtige Schätzung gewählt, da angenommen wird, daß die vergleichbaren **Error! Bookmark not defined.** Sorten die unendliche Zahl möglicher vergleichbarer **Error! Bookmark not defined.** Sorten für die Art insgesamt genau darstellen. Deshalb ist 1,6 der Schwellenwert für fremdbefruchtende Arten mit 60 Messungen je Merkmale und Sorte. Für verschiedene Stichprobengrößen sollte für  $df_1$  eine verschiedene Kenngröße F benutzt werden, obwohl  $df_2$  bei  $\infty$  bleiben sollte.

### 10.2 Schwellenwert für verschiedene Stichprobengrößen

10.2.1 Verschiedene Schwellenwerte von F (bei  $P = 0,01$ ) sollten für verschiedene Stichprobengrößen der Kandidatensorte angewandt werden. Die  $df_1$  variieren gemäß den verschiedenen Stichprobengrößen der Kandidatensorte. In allen Fällen wird  $df_2$  jedoch als  $\infty$  angesehen, um die gesamte Spannweite der möglichen vergleichbaren **Error! Bookmark not defined.** Sorten innerhalb einer Art zu erfassen und somit eine vorsichtige Schätzung des Schwellenwerts vorzunehmen. Unter diesen Bedingungen und wenn die entsprechenden Werte aus der F-Tabelle genommen werden, zeigt die Tabelle 1 den Schwellenwert, daß Grenzen für verschiedene Stichprobengrößen der Kandidatensorten gelten würden. Im Fall anderer Stichprobengrößen als der in Tabelle 1 enthaltenen sollte der richtige Schwellenwert für die exakte Stichprobengröße angewandt werden.

*Tabelle 1: Schwellenwert für die relative Varianz für einige verschiedene Stichprobengrößen*

<b>Stichprobengröße der Kandidatensorte</b>	<b>Schwellenwert für die relative Varianz</b>
30	2,03
40	1,81
50	1,68
60	1,60
80	1,49
100	1,43
150	1,33
200	1,28

Quelle: Tabelle von F, veröffentlicht in 'Tables for Statisticians', Barnes & Noble, Inc., New York

10.2.2 Für eine gegebene Stichprobengröße wird die Kandidatensorte für dieses Merkmal als nicht homogen angesehen, wenn die relative Varianz den Schwellenwert übersteigt.

### **10.3 Der Test der relativen Varianz in der Praxis**

10.3.1 Wenn die berechnete relative Varianz geringer als der tabellierte Wert der Kenngröße F in Tabelle 1 für die entsprechende Stichprobengröße ist, dann ist es angemessen anzunehmen, daß die Varianzen gleich sind und die Kandidatensorte in diesem bestimmten Merkmal homogen ist. Wenn die berechnete relative Varianz höher als der tabellierte Wert von F ist, dann wird die Nullhypothese, daß die Sorten gleiche Varianzen haben, zurückgewiesen. Die Kandidatensorte würde dann so angesehen, daß sie eine höhere Varianz als die vergleichbaren Sorten für dieses bestimmte Merkmal hat und würde daher die Homogenitätskriterien nicht erfüllen.

### **10.4 Beispiele für das Verfahren der relativen Varianz**

#### *Beispiel 1*

10.4.1 In einem DUS-Anbauversuch wird eine fremdbefruchtende Kandidatensorte zusammen mit einer Anzahl von Sorten angebaut, die das erforderliche Homogenitätsniveau in allen maßgebenden Merkmalen repräsentieren. Um die Berechnung der relativen Varianz zu verdeutlichen, wird ein Beispiel mit vier vergleichbaren Sorten gegeben. Die Varianzdaten für die Messungen der Pflanzhöhe der fünf Sorten sind in Tabelle 2 aufgeführt. Für jede Sorte wurde bei 60 Pflanzen die Pflanzhöhe gemessen:

*Tabelle 2: Varianzen der Kandidatensorte und der vergleichbaren Sorten für Pflanzhöhedaten*

Kandidatensorte	vergleichbare Sorte 1	vergleichbare Sorte 2	vergleichbare Sorte 3	vergleichbare Sorte 4
5,6	7,8	4,5	3,2	5,8

10.4.2 Die Zahl der Erfassungen je Sorte ist gleich ( $n=60$ ); deshalb kann die durchschnittliche Varianz der vergleichbaren Sorten als ihre zusammengefaßte Varianz genommen werden.

10.4.3 Die durchschnittliche Varianz für die vergleichbaren Sorten ist  $(7,8 + 4,5 + 3,2 + 5,8)/4 = 5,32$ .

10.4.4 Die relative Varianz für ein bestimmtes Merkmal bezieht sich auf die Varianz der Kandidatensorte, dividiert durch den Durchschnitt der Varianz der vergleichbaren Sorten.

Relative Varianz = Varianz der Kandidatensorte/durchschnittliche Varianz der vergleichbaren Sorten

$$= 5,6/5,32 = 1,05$$

10.4.5 Nun ist in Tabelle 1 für eine Stichprobengröße von 60 der Schwellenwert 1,60; deshalb kann gefolgert werden, daß die Kandidatensorte für dieses Merkmal hinreichend homogen ist.

## **10.5 Beziehung zwischen relativer Varianz und relativer Standardabweichung**

10.5.1 Bei DUS-Anbauversuchen werden die Homogenitätsdaten mitunter als Standardabweichungen, nicht als Varianzen angegeben. Mathematisch gesehen, besteht folgende einfache Beziehung zwischen Varianz und Standardabweichung:

Standardabweichung = Quadratwurzel der Varianz

10.5.2 Wenn wir mit relativen Standardabweichungen zu tun haben, muß daher die Tabelle 1 geändert werden, um die Quadratwurzeln der in Tabelle 4 angegebenen Schwellenwerte einzubeziehen.

*Tabelle 4: Schwellenwert für relative Standardabweichungen für einige verschiedene Stichprobengrößen*

<b>Stichprobengröße der Kandidatensorte</b>	<b>Schwellenwert für relative Standardabweichungen</b>
30	1,42
40	1,35
50	1,30
60	1,26
80	1,22
100	1,20
150	1,15
200	1,13

10.5.3 Bei der Entscheidung über die Homogenität aufgrund relativer Standardabweichungen muß der Prüfer die Tabelle 4 anstelle der Tabelle 1 benutzen, um die geeigneten Schwellenwerte zu erhalten. Derselbe Grundsatz für die Akzeptanz oder Zurückweisung gilt für die relative Standardabweichung; lediglich die Schwellenwerte sind infolge der Quadratwurzel der geeigneten Werte niedriger. Für 60 Stichproben beträgt der Schwellenwert der relativen Varianz beispielsweise 1,60; für die relative Standardabweichung beträgt der Schwellenwert jedoch 1,26, was die Quadratwurzel von 1,60 ist.

[Ende des Dokuments]